

AWARD NUMBER: W81XWH-13-1-0451

TITLE: The Genomic Evolution of Prostate Cancer

PRINCIPAL INVESTIGATOR: David J. VanderWeele, M.D., Ph.D.

RECIPIENT: The Geneva Foundation
Tacoma WA 98402

REPORT DATE: June 2017

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE June 2017			REPORT TYPE Annual		3. DATES COVERED 15 May 2016 - 14 May 2017	
4. TITLE AND SUBTITLE The Genomic Evolution of Prostate Cancer					5a. CONTRACT NUMBER	
					5b. GRANT NUMBER W81XWH-13-1-0451	
					5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Dr. David VanderWeele E-Mail: david.vanderweele@nih.gov; djvanderweele@gmail.com					5d. PROJECT NUMBER	
					5e. TASK NUMBER	
					5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) The Geneva Foundation 917 Pacific Avenue, Suite 600 Tacoma, Washington 98402					8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012					10. SPONSOR/MONITOR'S ACRONYM(S) CDMRP	
					11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited						
13. SUPPLEMENTARY NOTES						
14. ABSTRACT The goal of this research proposal is to characterize the oncogenesis of prostate cancer and to more fully understand the relationship between clinically benign and aggressive disease. With help from my mentors I have identified four areas where I will focus my training during the completion of the proposed project: 1. To continue to acquire a comprehensive understanding of prostate cancer genomics. 2. To develop an understanding of evolutionary genomics and tumor evolution. 3. To develop skills in translational prostate cancer biomarker development 4. To continue to develop laboratory management and grant writing skills.						
15. SUBJECT TERMS Cancer genetics, tumor evolution, tumor heterogeneity, prostate cancer, exome sequencing						
16. SECURITY CLASSIFICATION OF:				17. LIMITATION OF ABSTRACT Unclassified	18. NUMBER OF PAGES 68	19a. NAME OF RESPONSIBLE PERSON USAMRAA
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified	19b. TELEPHONE NUMBER (include area code)			

Table of Contents

	<u>Page</u>
1. Introduction	4
2. Keywords	4
3. Overall Project Summary	4
4. Key Research Accomplishments	6
5. Conclusion	6
6. Publications, Abstracts, and Presentations	6
7. Inventions, Patents and Licenses	8
8. Reportable Outcomes	8
9. Other Achievements	8
10. Opportunities for Training	8
11. Opportunities for Professional Development	8
12. References	9
13. Appendices	9

1. INTRODUCTION:

Prostate cancer is heterogeneous, with diverse outcomes. It is unclear to what extent low grade, indolent cancer can evolve into aggressive disease. And in aggressive disease, it is unclear if the genetic alterations more common in late disease are present early on, but at low frequency, or if they only develop late in disease progression. This project aims to characterize the evolution of both localized and metastatic prostate cancer. One aim seeks to evaluate the evolution of localized prostate cancer to high grade or aggressive variants. The second aim seeks to characterize the evolution from localized to metastatic prostate cancer.

2. KEYWORDS:

Cancer genetics, tumor evolution, tumor heterogeneity, prostate cancer, exome sequencing

3. OVERALL PROJECT SUMMARY BY SPECIFIC AIM:

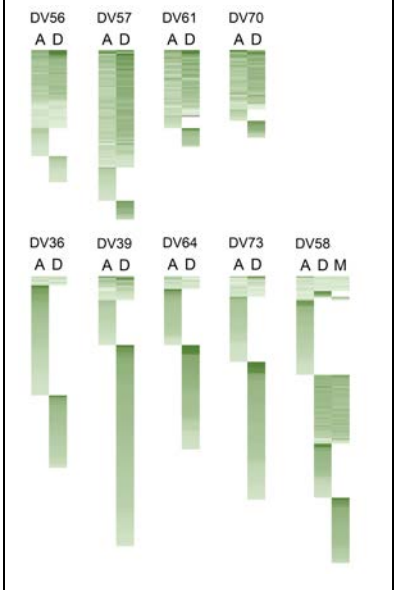
The research project proposed in this award was comprised of two specific aims, with further division into sub aims, and tasks outlined in the statement of work. This portion of the report will give a brief overall summary, followed by specific reference to the tasks outlined in the SOW.

Aim 1: Investigate the molecular relationship between coincident prostate cancer foci.

We set out to characterize the molecular relationship between coincident foci in two scenarios: coincident low and high grade disease; and coincident acinar and ductal histologies. In the first year we published that coincident high and low grade prostate cancer diverge early in their evolution, with only a minority of mutations being clonal (VanderWeele et al, Cancer Science, 2014), addressing the first question. We are now finalizing a manuscript detailing the relationship between acinar and ductal adenocarcinoma. We find that they are generally clonally related, indicating they arise from the same cell of origin (**Figure 1**). However, there are specific alterations that are found much more frequently in ductal foci than acinar foci and likely contribute to the histologic appearance.

While addressing this aim it was noted that mutations identified in metastases were found at low frequency in high grade, localized foci in the prostate. In collaboration with prostate cancer geneticists at RIKEN and University of Tokyo we characterized the genomic heterogeneity within individual index foci of aggressive, localized or locally advanced, prostate cancer. Using a novel tissue banking method (Gillard et al, American Journal of Translational Research, 2015) we performed multiregion genomic analysis for 10 cases of potentially lethal prostate cancer. Our results demonstrate marked intratumoral genetic heterogeneity. In subclonal populations we identify alterations that are more common in advanced, castrate resistant prostate cancer. Moreover, there is heterogeneity in genomic biomarkers which predict response to targeted therapies (**Figure 2**). The manuscript from this work is currently under review.

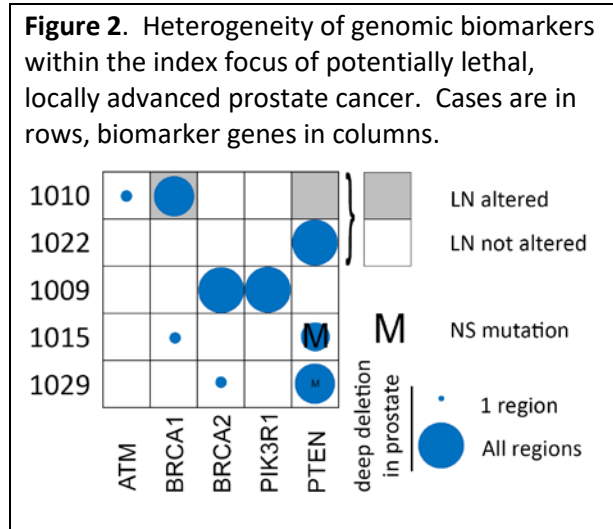
Figure 1. Mutations in coincident ductal and acinar adenocarcinoma of the prostate. Mutations are represented by rows. Darker green represents higher variant allele frequency. A=acinar, D=ductal, M=metastasis.



Aim 2: Characterize the molecular relationship between metastatic disease and primary prostate foci.

This aim evaluated the relationship between metastatic and primary disease in two scenarios. The first was primary prostate cancer and nodal metastases, where we found a close relationship between high grade disease and metastases (VanderWeele et al, Cancer Science, 2014), and in the manuscript currently under review.

The second scenario was comparing treatment-naïve disease, late stage CRPC, and circulating tumor cells (CTCs). Here we found that there is progression of the tumor genome, and this is captured by the genomics of the CTCs (Lack et al, Journal of Translational Medicine, 2017).



PROGRESS REPORT BY SOW TASKS

Task 1: Obtain tissue from matched prostate cancer foci and prepare sequencing libraries

- 1.1 Completed, manuscript in preparation
- 1.2 Completed, manuscript published (Lack et al, JTM, 2017)
- 1.3 Completed, manuscript in preparation
- 1.4 Completed, manuscript published (Lack et al, JTM, 2017)
- 1.5 Completed, manuscript under review
- 1.6 In progress

Task 2. Generate next-generation sequencing data on Illumina platform with massively parallel short (100 bp) reads, confirm mutations with Sanger sequencing or targeted resequencing

- 2.1 Completed, manuscript published (VanderWeele et al, Cancer Science, 2014)
- 2.2 Completed, manuscript published (Lack et al, JTM, 2017)
- 2.3 Completed, manuscript published (VanderWeele et al, Cancer Science, 2014)
- 2.4 Completed, manuscript published (Lack et al, JTM, 2017) and manuscript in preparation
- 2.5 Completed, manuscript under review
- 2.6 Completed, manuscript under review
- 2.7 Not yet completed

Task 3. Alignment of sequencing reads, calling mutations, and constructing phylogeny

- 3.1 Completed, manuscript published (VanderWeele et al, Cancer Science, 2014)
- 3.2 Completed, manuscript published (Lack et al, JTM, 2017)
- 3.3 Completed, manuscript published (Lack et al, JTM, 2017) and manuscript in preparation
- 3.4 Completed, manuscript published (Lack et al, JTM, 2017) and manuscript in preparation
- 3.5 Not yet completed

Mentorship

Dr. Kathy Kelly is my primary mentor, and she has continued to promote my professional development. I meet with her individually and with my fellow investigators in our department on a monthly basis. The most recent directions of my research have been strongly positively influenced by her work and the resources available within our department.

4. KEY RESEARCH ACCOMPLISHMENTS FOR YEAR 3:

- Acceptance of 2 manuscripts, including one on evolution of the genome from untreated disease to metastatic disease and circulating tumor cells (Aim 2b)
- Young Investigator Award from the Prostate Cancer Foundation

5. CONCLUSION:

Prostate cancer has a widely varying prognosis, with widely varying underlying genomics. The aggressive variant ductal adenocarcinoma, however, does not appear to have distinctive molecular features that distinguish it from conventional acinar adenocarcinoma, including similar level of Androgen Receptor activity. The similar genomic features would suggest that it would respond to standard therapies in a manner similar to aggressive conventional prostate cancer.

High risk prostate cancer, even when untreated, harbors alterations similar to that of advanced, castrate resistant disease. These alterations include biomarkers predictive of response to targeted therapy. Testing for these biomarkers should be undertaken with caution, however, since they are often subclonal. This implies that the results of biomarker tests would depend on where the tissue is sampled, and the indicated therapy may not be effective against all tumor cells.

Manuscripts are being prepared for submission to report on the findings above. Future lab work will focus on *in vitro* organoid models and *in vivo* xenograft models to characterize the evolution of tumor cells in response to therapy.

6. PUBLICATIONS, ABSTRACTS, AND PRESENTATIONS:

a. Peer-reviewed Scientific Journals

Year 1

1. VanderWeele DJ, Brown CB, Taxy JB, Stadler WM, White KP. Low grade prostate cancer foci diverge early from high grade and metastatic disease. *Cancer Science*. 2014 Aug;105(8):1079-85.
2. Reyes E*, VanderWeele DJ*, Duggan R, Campanile A, Stadler WM, Vander Griend DJ, Szmulewitz R. Quantitative characterization of androgen receptor expression in circulating tumor cells from patients with metastatic castrate resistant prostate cancer. *Journal of Translational Medicine*. 2014 Nov 26;12(1):313.

Year 2

1. Packiam V, Patel SG, Pariser JJ, Richards KA, Weiner AB, Paner GP, VanderWeele DJ, Zagaja GP, Eggener SE. "Contemporary Population-Based Comparison of Localized Ductal Adenocarcinoma and High-Risk Acinar Adenocarcinoma of the Prostate". *Urology*. 2015 Jul 18.
2. VanderWeele DJ, Paner GP, Fleming GF, Szmulewitz RZ. Sustained complete response to cytotoxic therapy and the PARP inhibitor veliparib in metastatic castration resistant prostate cancer – a case report. *Frontiers in Oncology*. 2015 5:169.
3. Gillard M, Tom W, Antic T, Paner GP, Lingen M, VanderWeele DJ. "Next-gen tissue: preservation of molecular and morphological fidelity in prostatectomy specimens". *American Journal of Translational Research*. 2015 Jul 15; 7(7):1227-1235.
4. McCann S, Jiang Y, Fan X, Wang J, Antic T, VanderWeele DJ, Oto A. "Quantitative Multi-parametric MRI Features and PTEN Expression in Prostate Cancer". *American Journal of Roentgenology*. 2016 Mar;206(3):559-65.

Year 3

1. Lack J, Gillard M, Cam M, Paner GP, VanderWeele DJ. "Circulating tumor cells capture disease evolution in advanced prostate cancer". *Journal of Translational Medicine*. 2017 Feb 23;15(1)44

b. Reviews and Book Chapters

Year 2

1. VanderWeele DJ, Stadler WM. "The time for molecular phenotyping of metastatic castrate-resistant prostate cancer is now". *Clinical Advances in Hematology & Oncology*. 2015 13(9), p. 580.

Year 3

1. VanderWeele DJ, Turkbey B, Sowalsky A. "Precision management of localized prostate cancer". *Expert Review of Precision Medicine and Drug Development*. 2016 Dec;1(6):505-15.
2. VanderWeele DJ, McNERNEY M, LeBeau M. "Chromosomal Aberrations in Cancer". *Holland-Frei Cancer Medicine*, 9th edition. Bast RC, Croce CM, Hait WN, et al, editors.

c. Presentations

Year 3

1. VanderWeele DJ, Finney R, Katayama K, Gillard M, Paner G, Imoto S, Yamaguchi R, Wheeler D, Cam M, Maejima K, Sasaki-Oku A, Nakano K, Tanaka H, Pontier A, Grigoryev D, Kubo M, Ratain M, Miyano S, Nakagawa H. "Local Enrichment of Multiple Clinically Significant Genetic Alterations within Index Foci of Locally Advanced Prostate Cancer." DOD IMPaCT meeting. Aug., 2016. Oral presentation.
2. VanderWeele DJ, Finney R, Katayama K, Gillard M, Paner G, Imoto S, Yamaguchi R, Wheeler D, Cam M, Maejima K, Sasaki-Oku A, Nakano K, Tanaka H, Pontier A, Grigoryev D, Kubo M, Ratain M, Miyano S, Nakagawa H. "Treatment-naïve, locally advanced prostate cancer contains numerous subclonal clinically significant alterations." Prostate Cancer Foundation annual scientific retreat. Oct., 2016. Poster presentation.

3. Vander Weele DJ, Lack J, Gillard M, Cam M, Paner GP. “Genomic analysis of untreated prostate adenocarcinoma, treatment-emergent neuroendocrine carcinoma, and circulating tumor cells (CTCs).” Genitourinary Cancer Symposium. Feb., 2017. Poster presentation.

7. INVENTIONS, PATENTS AND LICENSES:

Nothing to report

8. REPORTABLE OUTCOMES:

Nothing to report

9. OTHER ACHIEVEMENTS:

The data from this work, and the protected time insured from this award, has led to the successful application for a Young Investigator Award from the Prostate Cancer Foundation.

10. OPPORTUNITIES FOR TRAINING:

As per the SOW, I have completed the following courses

- Responsible Conduct of Research
- ECEV 35600 Principles of Population Genetics I
- ECEV 35901 Evolutionary Genomics
- Fundamentals of Clinical Research
- HGEN 47400 Introduction to Probability and Statistics for Geneticists
- STAT 22000 Statistical Methods & Their Applications
- Introduction to the Principles and Practice of Clinical Research
- Ethical and Regulatory Aspects of Clinical Research

I have yet to complete

- Principles of Clinical Pharmacology

11. OPPORTUNITIES FOR PROFESSIONAL DEVELOPMENT:

I had the opportunity to attend and participate in international conferences during the third year of this award. I gave an oral presentation at the Department of Defense IMPaCT meeting in August, 2016. I received valuable feedback on the project after this presentation.

I gave poster presentations at the Prostate Cancer Foundation annual scientific retreat and at the GU Cancers Symposium (GU ASCO). The work presented at GU ASCO was published shortly thereafter, describing the relationship between treatment-naïve prostate cancer, metastatic disease, and CTCs. Both of these meetings have clinical and translational components.

Though not directly related to my work in the lab, I gave an oral presentation on prostate cancer genetics and its clinical implications at the International Prostate Imaging Symposium. The work supported by this award led to the recognition of my expertise in this area and the invitation to speak. Similarly, the expertise gained from this award supported my successful application for a Young Investigator Award from the Prostate Cancer Foundation.

Finally, I have established collaborations across diverse fields to conduct a pre-clinical trial evaluating a novel class of Androgen Receptor (AR) antagonist, and to conduct correlative studies in conjunction with a clinical trial of neoadjuvant AR-targeted therapy.

12. REFERENCES:

1. VanderWeele DJ, Brown CB, Taxy JB, Stadler WM, White KP. Low grade prostate cancer foci diverge early from high grade and metastatic disease. *Cancer Science*. 2014 Aug;105(8):1079-85.
2. Gillard M, Tom W, Antic T, Paner GP, Lingen M, VanderWeele DJ. "Next-gen tissue: preservation of molecular and morphological fidelity in prostatectomy specimens". *American Journal of Translational Research*. 2015 Jul 15; 7(7):1227-1235.
3. Lack J, Gillard M, Cam M, Paner GP, VanderWeele DJ. "Circulating tumor cells capture disease evolution in advanced prostate cancer". *Journal of Translational Medicine*. 2017 Feb 23;15(1)44.

13. APPENDICES:

1. VanderWeele DJ, Brown CB, Taxy JB, Stadler WM, White KP. Low grade prostate cancer foci diverge early from high grade and metastatic disease. *Cancer Science*. 2014 Aug;105(8):1079-85.
2. Gillard M, Tom W, Antic T, Paner GP, Lingen M, VanderWeele DJ. "Next-gen tissue: preservation of molecular and morphological fidelity in prostatectomy specimens". *American Journal of Translational Research*. 2015 Jul 15; 7(7):1227-1235.
3. Lack J, Gillard M, Cam M, Paner GP, VanderWeele DJ. "Circulating tumor cells capture disease evolution in advanced prostate cancer". *Journal of Translational Medicine*. 2017 Feb 23;15(1)44.
4. VanderWeele DJ et al, manuscript under review at JNCI.

Low-grade prostate cancer diverges early from high grade and metastatic disease

David J. VanderWeele,^{1,2} Christopher D. Brown,^{1,3,6} Jerome B. Taxy,^{4,7} Marc Gillard,² David M. Hatcher,⁵ Westin R. Tom,⁵ Walter M. Stadler² and Kevin P. White^{1,2,3}

¹Institute for Genomics and Systems Biology, Departments of ²Medicine, ³Human Genetics, ⁴Pathology, ⁵Surgery, University of Chicago, Chicago, Illinois, USA

Key words

Carcinoma/genetics, exomes, genetic heterogeneity, multiple primary neoplasms, prostatic neoplasms

Correspondence

David VanderWeele, Department of Medicine, Section of Hematology/Oncology, The University of Chicago, 5841 South Maryland Avenue MC2115, Chicago, Illinois 60637, USA.
Tel: 773-834-7193; Fax: 773-702-9268;
E-mail: dvanderw@medicine.bsd.uchicago.edu

⁶Present address: Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

⁷Present address: Department of Pathology and Laboratory Medicine, NorthShore University Health System, Evanston, Illinois, USA

Funding information

This study is supported by the AACI Fellowship for Translational Cancer Research and DOD Prostate Cancer Research Program PRTA (DVW) and the Chicago Cancer Genomes Project (KPW).

Received January 8, 2014; Revised May 16, 2014; Accepted May 23, 2014

Cancer Sci 105 (2014) 1079–1085

doi: 10.1111/cas.12460

Cancers are thought to progress from low grade to higher grade,⁽¹⁾ leading to efforts to detect and eradicate low-grade disease in order to prevent morbidity and mortality from high-grade disease. The grade of prostate cancer is characterized by the Gleason score on a scale of 2–10, with 10 representing the most poorly differentiated tumors. Although initially described as intermediate grade, Gleason 6 cancers are clinically indolent and have an overwhelmingly favorable prognosis.^(2,3) For the purposes of this report, Gleason 6 foci are regarded as low grade.

Approximately 238 590 men were diagnosed with prostate cancer in the US in 2013, the majority of them with low-grade disease. It is estimated that 29 720 men will die from prostate cancer, which is among the lowest case/fatality ratio of any cancer.⁽⁴⁾ Although prostate specific antigen (PSA) screening is able to identify early stage prostate cancer, most men diagnosed with prostate cancer die from competing causes. It is therefore unclear whether early detection associated with PSA screening decreases mortality.^(5,6) Current clinical treatment guidelines support active surveillance for patients with

Understanding the developmental relationship between indolent and aggressive tumors is central to understanding disease progression and making treatment decisions. For example, most men diagnosed with prostate cancer have clinically indolent disease and die from other causes. Overtreatment of prostate cancer remains a concern. Here we use laser microdissection followed by exome sequencing of low- and high-grade prostate cancer foci from four subjects, and metastatic disease from two of those subjects, to evaluate the molecular relationship of coincident cancer foci. Seventy of 79 (87%) high-confidence somatic mutations in low-grade disease were private to low-grade foci. In contrast, high-grade foci and metastases harbored many of the same mutations. In cases in which there was a metastatic focus, 15 of 80 (19%) high-confidence somatic mutations in high-grade foci were private. Seven of the 80 (9%) were shared with low-grade foci and 65 (82%) were shared with metastatic foci. Notably, mutations in cancer-associated genes and the p53 signaling pathway were found exclusively in high-grade foci and metastases. The pattern of mutations is consistent with early divergence between low- and high-grade foci and late divergence between high-grade foci and metastases. These data provide insights into the development of high-grade and metastatic prostate cancer.

low-grade prostate cancer for which the risk of tumor progression is low.⁽⁷⁾ Recently some guidelines have even recommended against PSA screening.⁽⁸⁾ Failure to follow these guidelines is likely due, in part, to concern that low-grade foci progress to higher grade and metastatic disease.

Several studies have contributed to a growing catalog of genetic variation in prostate cancer,^(9–14) giving insight into single nucleotide variations, rearrangements and copy number variation prevalent in this disease. Aside from common rearrangements involving the ETS family of transcription factors, especially v-ets avian erythroblastosis virus E26 oncogene homolog (ERG), the mutational spectrum of prostate cancers is diverse.⁽¹⁰⁾ Although increasing attention is being paid to the overall molecular heterogeneity of cancers within a single patient, there is limited data in this regard specifically for prostate cancer. This is especially relevant for prostate cancer, which is often multifocal.

More specifically, it is not clear if prostate cancer develops as a multifocal disease in which the foci are clonally related and share the vast majority of somatic mutations, or if it is a

multicentric disease, with each focus having its own set of private somatic mutations and each representing an independent primary lesion. Recent studies have examined multifocal disease and found evidence of a monoclonal origin of coincident foci,⁽¹⁵⁾ even when the foci have different Gleason patterns.^(16,17) This conclusion is somewhat surprising, given the difference in clinical behavior of different Gleason patterns.

The present study targets multifocal disease, specifically with low-grade (Gleason 6) and high-grade (Gleason 8 or higher) foci, and compares the molecular relationship of these foci with each other and with synchronous metastatic disease. To expand the number of specimens available for the present study, prostatectomy samples that had been formalin fixed and paraffin embedded (FFPE) were used, with modifications made to standard protocols to increase the yield of material used for sequencing. Exome sequencing was used to achieve a comprehensive evaluation of all genes harboring somatic point mutations. The findings indicate that there is early divergence or complete independence of low-grade and high-grade disease, but late divergence of high-grade disease and synchronous metastases.

Materials and Methods

Tissue and laser capture microdissection (LMD). Paraffin-embedded prostate and lymph node tissue that was 1–5 years old and matched FFPE and frozen normal prostate tissue was obtained from the Human Tissue Resource Center (HTRC) at the University of Chicago with the approval of the Internal Review Board. Tissue was reviewed by a genitourinary pathologist (J.B.T.) and regions of interest were marked for further evaluation. Four specimens had sufficient low- and high-grade disease for LMD. Two specimens (PrCa 6 and PrCa 18) also had sufficient lymph node metastatic disease. Tumor foci and histologically normal prostate were microdissected from 10- μ m sections using a Leica Laser Microdissection (Wetzlar, Germany) system.

Exome library preparation. Genomic FFPE DNA was isolated using a QIAamp DNA FFPE Tissue Kit (Qiagen, Valencia, CA, USA) following the manufacturer's instructions, but increasing the incubation in Proteinase K to overnight and reducing the 90°C incubation to 30 min. DNA from frozen tissue was isolated using a QIAamp Mini Kit (Qiagen). DNA

was fragmented using a Covaris S2 sonicator (Woburn, MA, USA), with settings of 10% duty cycle, intensity 4, 200 cycles per burst and 120 s. The fragmented DNA was blunt-end repaired using T4 Polymerase, Klenow and T4 PNK, and adenylated. Fragments were ligated to adaptors IS1_adapter_P5.F and IS3_adapter_P5 + P7.R⁽¹⁸⁾ and an index tag added. The product was amplified and cleaned up using Agencourt AM-Pure XP beads (Beckman Coulter, Indianapolis, IN, USA).

Libraries were run on a BioAnalyzer (Agilent, Santa Clara, CA, USA) to confirm fragment size. The libraries from the metastases did not undergo further size selection. The remaining libraries underwent selection using an E-gel (Invitrogen, Grand Island, NY, USA) followed by PCR amplification.

Enrichment for exomes was performed using NimbleGen SeqCap EZExome v2.0 (Roche, Madison, WI, USA) following the manufacturer's instructions. The exome-enriched library DNA was amplified and run on a BioAnalyzer to confirm fragment size. The resulting libraries were run on a HiSeq 2000 (Illumina, San Diego, CA, USA) in a highly multiplexed 1 \times 50 run and a subsequent 2 \times 100 run.

Sequencing analysis and identification of somatic mutations. Overlapping or adapter-containing reads were clipped using SeqPrep-b83fd00. Reads were quality trimmed and aligned to hg18 using bwa 0.5.9⁽¹⁹⁾ and merged with duplicates removed using Samtools 0.1.18.⁽²⁰⁾ Genome Analysis Toolkit⁽²¹⁾ was used for local realignment and quality score recalibration. Somatic mutations were identified using VarScan v2.2.8.⁽²²⁾ High-confidence somatic mutations were defined as variants with coverage of at least 10 \times , no supporting reads in normal, at least four supporting reads in the tumor and with supporting reads on both strands. Germline single-nucleotide polymorphisms (SNPs) identified in dbSNP135, 1000 genomes and exomes of normal controls published by the Max Planck Institute (MPI)^(23–25) were removed. SAMtools was used to determine the number of total and variant reads in matched libraries for positions mutated in low-grade foci. Sample ethnicity was determined by principle component analysis using common, germline, coding variants shared with samples from the 1000 genomes project.⁽²⁴⁾

Custom capture and identification of previously identified mutations. DNA from seven additional foci from PrCa 6 were isolated using LMD (Leica Microsystems LMD6500). Custom capture probes targeting high-confidence mutations

Table 1. Subject characteristics

Case	Patient age (years)	Race	Stage	Gleason – gland†	Library	Gleason – focus‡
PrCa 18	65	Caucasian	pT3bN1Mx	4 + 3	Normal	NA
					Low grade	6
					High grade	8
					Metastasis	NA
PrCa 14	54	African–American	pT3bN1Mx	4 + 3	Normal	NA
					Low grade	6
					High grade	8
PrCa 6	70	Caucasian	pT3bN1Mx	4 + 3 (5)	Normal	NA
					Low grade	6
					High grade	8
					Metastasis	NA
PrCa 25	59	Asian	pT2cN0Mx	3 + 4	Normal	NA
					Low grade	6
					High grade	9

†Gleason score assigned on pathological review of the prostatectomy specimen. ‡Gleason score of the focus isolated by microdissection. PrCa 6 had a Gleason score of 4 + 3 with a minor component of 5. PrCa 14 had metastatic disease in a lymph node, which was not of sufficient mass to be isolated for sequencing. NA, not applicable.

identified through exome sequencing were obtained from Roche. Sequencing libraries from the seven loci were prepared using the High-Throughput Genome Analysis Core at the University of Chicago, pooled, captured following the manufacturer's instructions, and sequenced on one lane of a HiSeq2000 (Illumina). Reads were analyzed with SeqPrep-b83fd00, aligned to hg18 using bwa, and merged with duplicates removed using Samtools 0.1.18. Samtools was used to determine the number of total and variant reads in each library for each variant position. Variants were considered present if >10% of reads supported the variant.

Confirmation of somatic mutations and immunohistochemistry (IHC). Mutations were visualized using IGV 2.0.⁽²⁶⁾ Genomic DNA from each focus underwent amplification using GenomePlex WGA2 (Sigma, St. Louis, MO, USA) and primers targeting a subset of variants were designed using Primer3Plus and obtained from IDT (Coralville, IA, USA). The PCR products were sequenced in the University of Chicago DNA Sequencing and Genotyping Facility. ERG IHC was performed at the

HTRC using anti-ERG antibody EPR3864 (Abcam, Cambridge, MA, USA).

Results

To evaluate the molecular relationship between low- and high-grade prostate cancer, we identified multifocal formalin-fixed, paraffin-embedded specimens with both low-grade and high-grade disease with sufficient mass for laser capture microdissection. The subjects from whom the specimens came were representative of patients with high-risk localized prostate cancer (Table 1).

Exome sequencing was performed on matched cancer foci and histologically normal prostate glands from four specimens (Fig. 1). To improve the yield of DNA for downstream sequencing, proteinase K digestion was increased to 18 h, 90° incubation was reduced to 30 min and size selection by gel extraction was replaced with size selection with an E-gel or eliminated altogether. On average, 88% of RefSeq coding

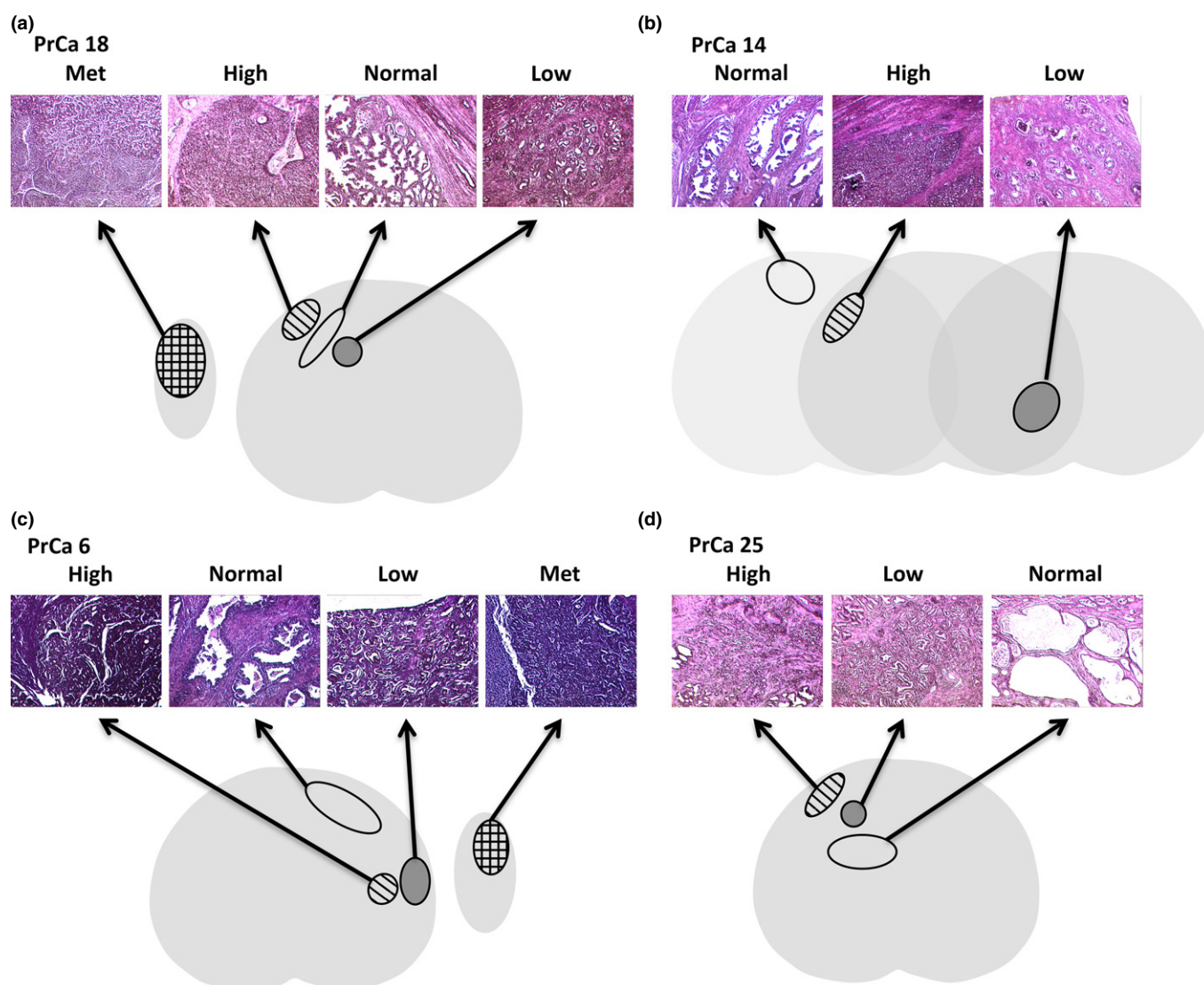


Fig. 1. Histology of coincident prostate cancer foci. Representative H&E stains and illustrations representing the prostate cancer foci that were laser microdissected. Two cancer foci and uninvolved prostate glands were isolated from PrCa 18 (a), PrCa 14 (b), PrCa 6 (c) and PrCa 25 (d). In addition, a metastatic (Met) focus was isolated from PrCa 18 and PrCa 6. For PrCa 14 (b), the foci were from different levels of the prostate. For the other three specimens the foci were at the same level. Light gray, histologically normal prostate; dark gray, low-grade cancer focus; striped, high-grade cancer focus; checked, metastatic focus from a lymph node removed at the time of prostatectomy.

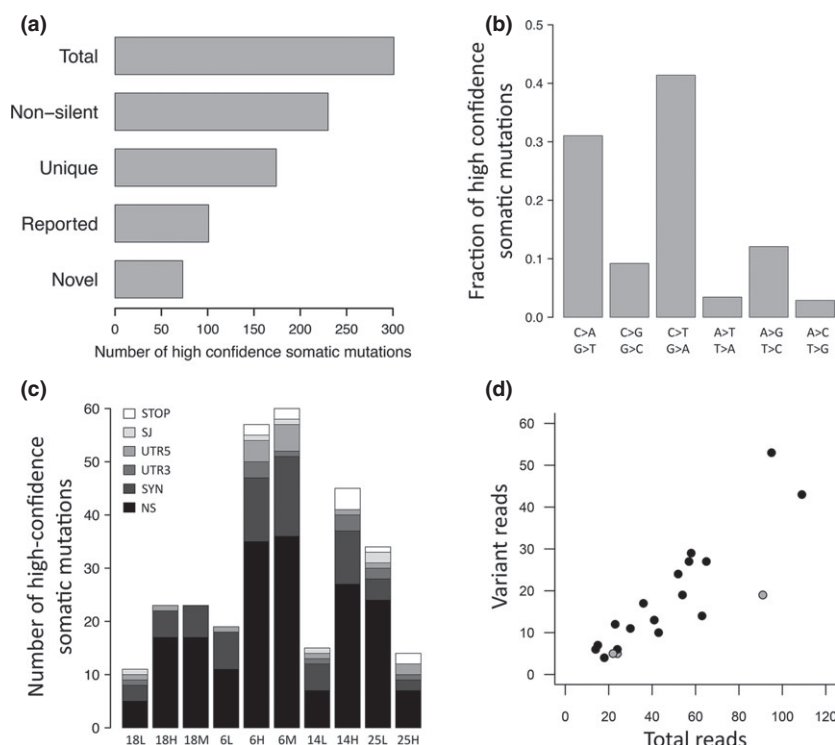


Fig. 2. Mutation characteristics. (a) Number of high-confidence somatic mutations across all foci. Non-silent, non-silent mutations; Unique, number of unique genes harboring a non-silent mutation; Reported, gene reported to be mutated in references 9–12 and 14. (b) Spectrum of unique high confidence somatic mutations across all foci. (c) Number of high-confidence somatic mutation types in all prostate samples. L, low-grade focus; H, high-grade focus; M, metastatic focus; NS, nonsynonymous; SYN, synonymous; UTR3, 3' untranslated region; UTR5, 5' untranslated region; SJ, splice junction; STOP, premature stop codon. (d) Confirmation of high-confidence somatic mutations with capillary sequencing by number of total and variant reads (gray, not confirmed; black, confirmed). Confirmation rate, 0.85.

bases were covered at 10× or greater. A mean of 30.1 somatic coding mutations was identified per cancer focus (Supporting Information Table S1).

To evaluate potential bias introduced by formalin fixation, matched FFPE and frozen (Optimal Cutting Temperature compound [OCT]-embedded) samples of normal prostate were subjected to exome sequencing. The number of false positive variant reads was similar (Fig. S1A), as were the types of mutations (Fig. S1B), suggesting that there was little bias introduced from formalin fixation.

Three hundred and one somatic mutations were identified across 10 foci (Table S2). Of these, 71 were synonymous and unlikely to alter gene function (Fig. 2a). Of 174 unique genes harboring somatic mutations that were likely to be functional, 73 were not reported in five previous reports of mutations in localized and advanced disease.^(9–12,14) The majority of variants identified were transitions leading to nonsynonymous mutations (Fig. 2b,c), as expected. A subset of 20 high-confidence somatic mutations was confirmed in tumor and uninvolved tissue with capillary sequencing (Fig. 2d). Those that were not confirmed by capillary sequencing had low variant allele frequency and were approximately the level of detection for capillary sequencing (~20%). Nevertheless, the mutation (0.82/Mb) and validation rate (85%) using fixed prostate specimens in the present study are similar to previous reports (0.9/Mb and 91%, respectively) using frozen tissue.⁽⁹⁾ The high fidelity of this sequencing supports the use of formalin-fixed tissue for next-generation sequencing studies.

The vast majority of somatic mutations in low- and high-grade foci were private to the focus in which they were identified. Nine of 79 (11%) high-confidence somatic mutations in the low-grade foci were shared with a high-grade focus, and nine of 139 (7%) high-confidence somatic mutations in the high-grade foci were shared with a low-grade focus (Fig. 3). Capillary sequencing confirmed that these nine shared somatic mutations were not germline variants. In two subjects (14 and 18) all of the mutations were private. In the other two subjects (6 and 25) a small

minority of mutations was shared. In contrast to the low-grade foci, 67 of 80 (84%) high-confidence somatic mutations identified in high-grade foci were shared with metastatic disease and metastatic foci had few mutations beyond those identified in high-grade disease (Fig. 3a,c).

To evaluate more broadly the clonal relationship of multifocal disease, custom capture baits targeting variants identified through exome sequencing of PrCa 6 were used to capture the DNA from seven additional cancer foci from that subject (Table S3). The average coverage depth of these variants was 208 and 78 previously identified variants were covered at >70× in each focus. Three high-grade foci harbored a median of 56 previously identified variants, three intermediate-grade foci harbored a median of 15 variants and one low-grade focus harbored three variants. As in the whole exome data a small minority of mutations was shared, with just two of 78 (2.6%) variants found to be ubiquitous among the seven cancer foci. Fourteen of 78 (17.9%) were not identified in any of the seven additional foci (Fig. S2). Eight of these 14 were among the 12 variants originally identified as private to the low-grade focus.

Rearrangements involving ERG are the most common genetic derangement in prostate cancer and are thought to occur early in tumorigenesis. These events are not typically captured by exome sequencing. Therefore we performed IHC for ERG to look for overexpression, which is typically restricted to cells harboring an ERG rearrangement. Of the two specimens exhibiting positive ERG staining, PrCa 6 showed consistently positive staining in all tumor foci (Fig. 4). In contrast, in PrCa 18 the low-grade focus was negative, the high-grade focus was equivocal with few cells positive for ERG and the metastatic focus was more widely focally positive.

Together with the immunohistochemical analysis, the patterns of shared and private mutations indicate early divergence of low- and high-grade foci. In two cases no shared mutations were identified and independent origin of low- and high-grade disease cannot be ruled out. In contrast, there is late divergence of high-grade foci and metastatic disease. Consistent

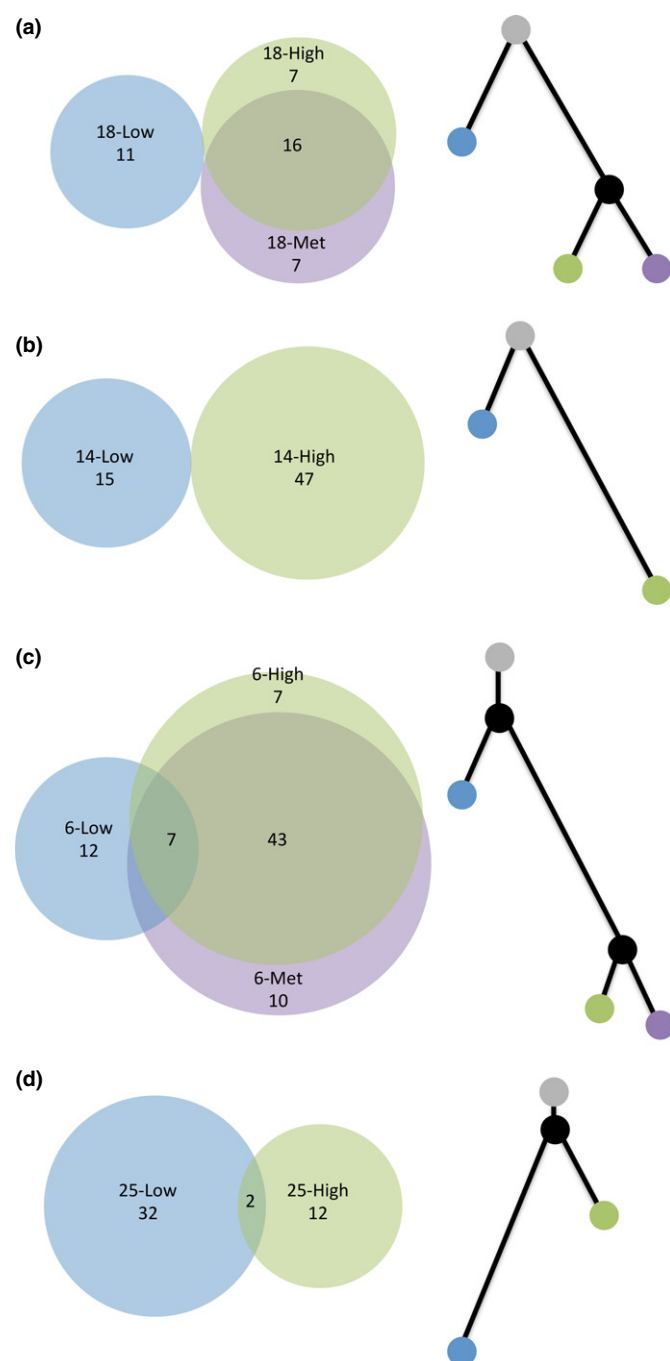


Fig. 3. The molecular relationship of coincident foci. Venn diagrams (left) depicting the pattern of shared and private high-confidence somatic mutations and phylogenetic trees (right) depicting the relationship of coincident foci for PrCa 18 (a), PrCa 14 (b), PrCa 6 (c) and PrCa 25 (d). The number of high-confidence somatic mutations is labeled within each Venn diagram. In (c), there are no mutations that are shared between the low-grade focus and the metastatic focus and also not shared with the high-grade focus. Within each phylogenetic tree, branch length is proportional to the number of mutations. Gray, uninvolved prostate; blue, low-grade focus; green, high-grade focus; purple, metastatic focus; black, theoretical common progenitor.

with this, there were no somatic mutations that were shared between low grade and metastatic disease and also not shared with high-grade disease.

Prostate tumor foci consist of an admixture of tumor cells and stromal cells. Despite microdissection of the tumor foci, the var-

iant allele frequencies in the low-grade foci were modestly lower than the high-grade lesions (Fig. S3), suggesting higher contamination with non-tumor cells. The contaminating stromal cells may interfere with the ability to call somatic mutations, especially in the low-grade foci, leading to a falsely elevated number of apparently private mutations and obscuring the clonal relationship among concurrent foci. Therefore for each position identified as a private high-confidence somatic mutation in the low-grade foci, the number of sequencing reads supporting that mutation in the matched high grade and metastatic foci was determined, relaxing filters based on DNA strand and quality scores (Fig. S4). Of the 23 high-confidence somatic mutations in PrCa 18 and PrCa 6 identified as private mutations, none had two or more variant reads in high grade or metastatic foci. Thus, even with a less conservative definition of mutation (two variant reads), high-grade foci and metastases harbored few variants seen in low-grade foci and the majority of variants in the low-grade foci was private. This was corroborated with evidence from capillary sequencing. Three mutations private to low-grade foci and two shared with other foci were among the subset of mutations confirmed by capillary sequencing.

Two genes exhibited recurrent nonsynonymous mutations or mutations in the untranslated region (Table S4). Low density lipoprotein (LDL) receptor-related protein 1B, which encodes an LDL-family receptor, is frequently deleted in multiple malignancies.⁽²⁷⁾ However, recent work suggests its high frequency of mutation may not imply a significant role in cancer.⁽²⁸⁾ ZNF717 is one of a number of zinc finger proteins and its role in malignancy is unclear.

Comparison of genes with nonsynonymous mutations in our data to the Catalogue Of Somatic Mutations In Cancer (COSMIC) cancer gene census revealed three genes in common (Table 2). These genes include TP53, the most frequently mutated gene in advanced cancers including prostate cancer,⁽¹²⁾ as well as ATM and SS18L1. All of these three genes were mutated in high grade or metastatic disease, but not in low-grade disease. Gene ontology analysis⁽²⁹⁾ revealed enrichment for genes in the p53 signaling pathway among high-grade foci (adjusted *P*-value 0.037), including TP53 and ATM. Cell cycle genes were also enriched (adjusted *P*-value 0.054). No other pathways were statistically significantly enriched.

Discussion

Given the disparate outcomes between men with low-grade prostate cancer and those with high-grade disease, a fundamental question in prostate cancer management is the possible molecular relationship among grades and metastatic disease. Multiple lines of evidence suggest multifocal disease is often independent.^(30,31) However, recent reports have demonstrated shared molecular derangements of coincident low- and high-grade foci.^(16,17) The present study identifies mutations shared between low- and high-grade cancer foci, supporting the ability of a single progenitor to give rise to both low- and high-grade disease. However, the relationship between these foci was distant relative to the relationship between high-grade disease and synchronous metastases. In fact, in two of the four cases it cannot be ruled out that the foci arose from independent origins. The early divergence of low- and high-grade disease may help explain distinct differences in their clinical behavior. Indeed, we found no evidence of direct progression from low grade to metastatic disease. Instead, there was an overwhelming fraction of shared mutations between high grade and metastatic disease, which is consistent with late divergence of these high-grade foci (Fig. 5).

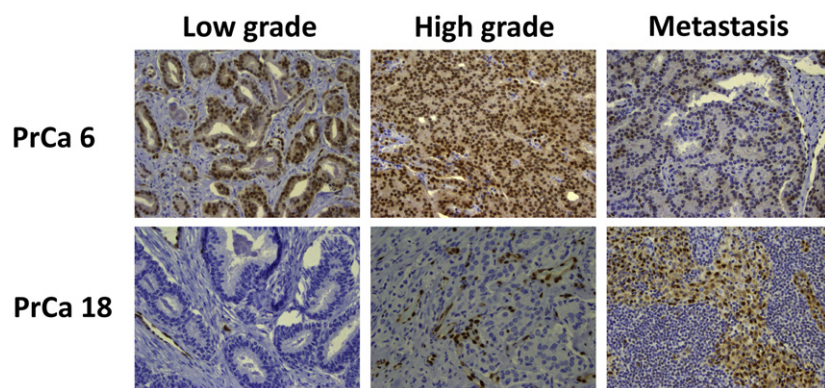


Fig. 4. ERG expression. Immunohistochemical analysis for ERG in coincident foci from PrCa 6 and PrCa 18.

Table 2. Overlap of mutated genes with the COSMIC cancer gene census

Subject	Histology	Chromosome	Hg18 position	Reference	Variant	Symbol	GERP	Type	aa Change
6	High	11	107675775	G	A	ATM	5.22	STOP	W1710*
6	LN	11	107675775	G	A	ATM	5.22	STOP	W1710*
6	High	20	60169915	C	G	SS18L1	3.98	NS	P87R
6	LN	20	60169915	C	G	SS18L1	3.98	NS	P87R
14	High	17	7517833	C	A	TP53	4.78	NS	C145F

Three genes harboring a nonsynonymous mutation or a premature stop codon are also found in the COSMIC cancer gene census. All three were found in either high-grade foci or metastatic disease, but not in low-grade foci. GERP, genomic evolutionary rate profiling score; LN, lymph node metastasis; NS, nonsynonymous; STOP, introduction of premature stop codon; *, stop codon.

For the cases in which there was evidence of a common progenitor for both low- and high-grade disease (PrCa 25 and PrCa 6), the nature of the common progenitor is not clear. Histological low-grade cancer, HGPIN or even histologically uninvolved prostate are all possible progenitors. If the progenitor is not histological low-grade cancer, high-grade and low-grade disease may have emerged simultaneously. Characterization of the common progenitor requires further study.

The present study demonstrates there is late divergence of high-grade cancer and synchronous metastatic disease, confirming high-grade disease as the precursor to metastasis and suggesting few if any additional mutations are required for malignant cells to gain the ability to metastasize. Given the difficulty in obtaining tissue from typical bone metastases in prostate cancer, more easily accessible prostate lesions might offer a substitute for identifying genetic derangements leading to primary resistance of metastatic disease. However, mechanisms of secondary resistance are likely induced or enriched by the selective pressure of therapy and thus are likely found at low levels, if at all, in the primary lesion.

The most common gene fusion event in prostate cancer is the TMPRSS2-ERG fusion, leading to high levels of expression of ERG, which is otherwise expressed at very low levels in prostate cancer. We evaluated for the presence of this fusion through IHC for ERG and those findings supported the evidence from the exome sequencing data. A limitation of exome sequencing is the inability to identify additional chromosomal rearrangements. Although this is a limitation it was not required for the goal of the present study, which was to identify tumor lineages.

One goal of sequencing prostate tumors is to identify biomarkers that will supplement the prognosis given by clinical features alone. To that end we sought to identify mutated genes or pathways that were enriched in high grade and metastatic disease. TP53 is the most frequently mutated gene in advanced prostate cancer and we found the p53 signaling pathway to harbor mutations only in high grade and metastatic disease. Thus, members of the p53 signaling pathway are

intriguing candidates for potential biomarkers of aggressive disease. Recently, expression of a panel of cell cycle progression-associated genes has been correlated with outcomes of those with low- or intermediate-grade prostate cancer.⁽³²⁾ We also found enrichment for cell cycle-related genes among those mutated in high-grade disease, although this was of borderline statistical significance in this small cohort.

Given the good outcomes of low-grade prostate cancer, we hypothesized that mutations not found in low-grade disease are more likely to be drivers of progression and metastasis and therefore more important potential therapeutic targets. Indeed, the three genes identified here that are also found in the COSMIC cancer gene census were identified exclusively in high grade and metastatic foci. Of note, all three mutations are in highly conserved regions as evidenced by the high genomic evolutionary rate profiling score and thus are likely to be biologically significant. Identification of driver genes through this strategy has the potential to guide the development of future therapies.

The present study contributes to the growing body of work demonstrating that next-generation sequencing from microdissected FFPE prostatectomy specimens is feasible. Based on these techniques, we found that coincident low- and high-grade prostate cancer can emerge through either independent progression or early divergence from a common progenitor.

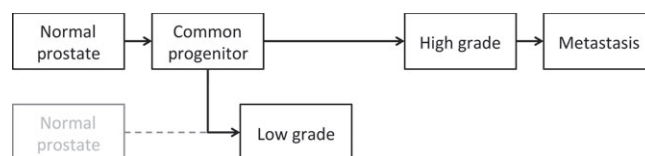


Fig. 5. Model of prostate cancer evolution. Low-grade and high-grade cancer foci progress largely in parallel, diverging early from a common progenitor. Metastatic disease exhibits late divergence from high-grade disease. It cannot be ruled out that in some cases low- and high-grade diseases arise through independent origins.

High-grade disease is characterized by mutations in known cancer-associated genes and the p53 signaling pathway. Metastatic disease is closely related to high-grade primary foci. These data can be leveraged to help identify potential biomarkers and drivers of progression of clinically significant disease.

Acknowledgements

The authors thank April Peterson for technical assistance in library preparation and A. Jason Grundstad, Dr Megan McNerney and Dr Thomas Stricker for advice on experimental design and data analysis.

The present study is supported by the AACI Fellowship for Translational Cancer Research and DOD Prostate Cancer Research Program PRTA (D.J.V.W.) and the Chicago Cancer Genomes Project (K.P.W.).

Disclosure Statement

The authors have no conflicts of interest.

References

- Vogelstein B, Kinzler KW. The multistep nature of cancer. *Trends Genet* 1993; **9**: 138–41.
- Egger SE, Scardino PT, Walsh PC *et al.* Predicting 15-year prostate cancer specific mortality after radical prostatectomy. *J Urol* 2011; **185**: 869–75.
- Ross HM, Kryvenko ON, Cowan JE, Wheeler TM, Epstein JI. Do adenocarcinomas of the prostate with Gleason score (GS) ≤ 6 have the potential to metastasize to lymph nodes? *Am J Surg Pathol* 2012; **36**: 1346–52.
- Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer J Clin* 2013; **63**: 11–30.
- Andriole GL, Crawford ED, Grubb RL 3rd *et al.* Mortality results from a randomized prostate-cancer screening trial. *N Engl J Med* 2009; **360**: 1310–9.
- Schröder FH, Hugosson J, Roobol MJ *et al.* Prostate-cancer mortality at 11 years of follow-up. *N Engl J Med* 2012; **366**: 981–90.
- Mohler J, Bahnson RR, Boston B *et al.* NCCN clinical practice guidelines in oncology: prostate cancer. *J Natl Compr Canc Netw* 2010; **8**: 162–200.
- Moyer VA. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2012; **157**: 120–34.
- Berger MF, Lawrence MS, Demicheli F *et al.* The genomic complexity of primary human prostate cancer. *Nature* 2011; **470**: 214–20.
- Barbieri CE, Baca SC, Lawrence MS *et al.* Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat Genet* 2012; **44**: 685–9.
- Grasso CS, Wu Y-M, Robinson DR *et al.* The mutational landscape of lethal castration-resistant prostate cancer. *Nature* 2012; **487**: 239–43.
- Kumar A, White TA, MacKenzie AP *et al.* Exome sequencing identifies a spectrum of mutation frequencies in advanced and lethal prostate cancers. *Proc Natl Acad Sci USA* 2011; **108**: 17087–92.
- Taylor BS, Schultz N, Hieronymus H *et al.* Integrative genomic profiling of human prostate cancer. *Cancer Cell* 2010; **18**: 11–22.
- Robbins CM, Tembe WA, Baker A *et al.* Copy number and targeted mutational analysis reveals novel somatic events in metastatic prostate tumors. *Genome Res* 2011; **21**: 47–55.
- Boyd LK, Mao X, Xue L *et al.* High-resolution genome-wide copy-number analysis suggests a monoclonal origin of multifocal prostate cancer. *Genes Chromosomes Cancer* 2012; **51**: 579–89.
- Sowalsky AG, Ye H, Bubley GJ, Balk SP. Clonal progression of prostate cancers from Gleason grade 3 to grade 4. *Cancer Res* 2013; **73**: 1050–5.
- Kovtun IV, Chevillat JC, Murphy SJ *et al.* Lineage relationship of Gleason patterns in Gleason score 7 prostate cancer. *Cancer Res* 2013; **73**: 3275–84.
- Meyer M, Kircher M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* 2010; **2010**: pdb.prot5448.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009; **25**: 1754–60.
- Li H, Handsaker B, Wysoker A *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* 2009; **25**: 2078–9.
- McKenna A, Hanna M, Banks E *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010; **20**: 1297–303.
- Koboldt DC, Zhang Q, Larson DE *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012; **22**: 568–76.
- Sherry ST, Ward MH, Kholodov M *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001; **29**: 308–11.
- Consortium T 1000 GP. A map of human genome variation from population-scale sequencing. *Nature* 2010; **467**: 1061–73.
- Burbano HA, Hodges E, Green RE *et al.* Targeted investigation of the Neanderthal genome by array-based sequence capture. *Science* 2010; **328**: 723–5.
- Robinson JT, Thorvaldsdóttir H, Winckler W *et al.* Integrative genomics viewer. *Nat Biotechnol* 2011; **29**: 24–6.
- Prazeres H, Torres J, Rodrigues F *et al.* Chromosomal, epigenetic and microRNA-mediated inactivation of LRP1B, a modulator of the extracellular environment of thyroid cancer cells. *Oncogene* 2011; **30**: 1302–17.
- Lawrence MS, Stojanov P, Polak P *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 2013; **499**: 214–8.
- Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2008; **4**: 44–57.
- Mehra R, Han B, Tomlins SA *et al.* Heterogeneity of TMPRSS2 gene rearrangements in multifocal prostate adenocarcinoma: molecular evidence for an independent group of diseases. *Cancer Res* 2007; **67**: 7991–5.
- Lindberg J, Klevebring D, Liu W *et al.* Exome sequencing of prostate cancer supports the hypothesis of independent tumour origins. *Eur Urol* 2013; **63**: 347–53.
- Cooperberg MR, Simko JP, Cowan JE *et al.* Validation of a cell-cycle progression gene panel to improve risk stratification in a contemporary prostatectomy cohort. *J Clin Oncol* 2013; **31**: 1428–34.

Supporting Information

Additional supporting information may be found in the online version of this article:

Fig. S1. Comparison of fixed and frozen.

Fig. S2. Variant allele frequency of variants in additional cancer foci.

Fig. S3. Variant allele frequencies.

Fig. S4. Frequencies of variants identified as private to the low-grade focus.

Table S1. Characteristics of sequencing libraries.

Table S2. High-confidence somatic mutations identified across all foci sequenced.

Table S3. Variants identified in additional cancer foci.

Table S4. Recurrently mutated genes.

Original Article

Next-gen tissue: preservation of molecular and morphological fidelity in prostate tissue

Marc Gillard¹, Westin R Tom², Tatjana Antic³, Gladell P Paner³, Mark W Lingen³, David J VanderWeele¹

Departments of ¹Medicine, ²Surgery, ³Pathology, University of Chicago, Chicago, IL

Received May 14, 2015; Accepted July 12, 2015; Epub July 15, 2015; Published July 30, 2015

Abstract: Background: Personalization of cancer therapy requires molecular evaluation of tumor tissue. Traditional tissue preservation involves formalin fixation, which degrades the quality of nucleic acids. Strategies to bank frozen prostate tissue can interfere with diagnostic studies. PAXgene is an alternative fixative that preserves protein and nucleic acid quality. Methods: Portions of prostates obtained from autopsy specimens were fixed in either 10% buffered formalin or PAXgene, and processed and embedded in paraffin. Additional sections were immediately embedded in OCT and frozen. DNA and RNA were extracted from the formalin-fixed, PAXgene-fixed, or frozen tissue. Quantitative PCR was used to compare the quality of DNA and RNA obtained from all three tissue types. In addition, 5 µm sections were cut from specimens devoid of cancer and from prostate cancer specimens obtained at prostatectomy and fixed in PAXgene. They were either stained with hematoxylin and eosin or interrogated with antibodies for p63, PSA and p504. Results: Comparable tissue morphology was observed in both the formalin and PAXgene-fixed specimens. Similarly, immunohistochemical expression of the P63, PSA and P504 proteins was comparable between formalin and PAXgene fixation techniques. DNA from the PAXgene-fixed tissue was of similar quality to that from frozen tissue. RNA was also amplified with up to 8-fold greater efficiency in the PAXgene fixed tissue compared to the formalin-fixed tissue. Conclusions: Prostate specimens fixed with PAXgene have preserved histologic morphology, stain appropriately, and have preserved quality of nucleic acids. PAXgene fixation facilitates the use of prostatectomy tissue for molecular biology techniques such as next-generation sequencing.

Keywords: Prostate cancer, PAXgene, formalin, fixation, next-generation sequencing, immunohistochemistry

Introduction

Prostate cancer (PCa) is the most common malignancy in many Western countries, affecting over 230,000 men each year in the U.S. alone, and it is the second leading cause of cancer death among American men [1]. Localized PCa is a clinically heterogeneous disease with significant variation in patient outcome. Determining the most beneficial management strategy requires risk stratification based on clinical factors and diagnostic/prognostic biomarkers [2]. Though advanced PCa is also clinically heterogeneous, most men are treated with the same therapies. Biomarker-driven therapies are just now beginning to enter into clinical trials, and none have entered into routine clinical practice.

It is essential to identify better predictors that could guide therapeutic actions and help reach

the target of personalized medicine in PCa. In this challenging context, next-generation sequencing (NGS) technologies constitute powerful tools to identify potential prognostic and predictive biomarkers based on each patient's mutational landscape and expression signature. The recent rise of sequencing of PCa genomes has significantly increased our understanding of the molecular basis of the disease [3-6]. Efforts are being made towards translating these new findings into clinical care [7, 8].

The development of biomarkers through NGS or other methods requires access to high quality specimens. In the case of PCa, this resource is particularly difficult to acquire. PCa is rarely visible on gross examination due to its small size or the intermingling of benign and malignant glands, making sampling difficult and often poorly representative of the tumor. Moreover, standard prostatectomy specimen processing

Preservation of prostate tissue in fidelity

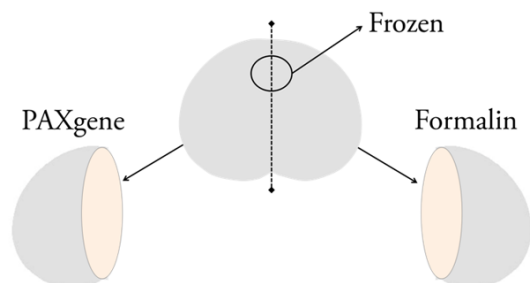


Figure 1. Autopsy prostate processing. Mirrored sections from four prostate specimens from autopsies were fixed either in PAXgene or formalin. Two specimens were fixed in PAXgene for 24 hours, and two specimens were fixed in PAXgene for 48 hours. For each prostate a frozen sample was taken before fixation.

requires that the gland should be fully fixed in formaldehyde solution before sectioning to guarantee the accurate assessment of surgical margins [9]. Others have proposed methods to freeze a larger fraction of the specimen, but these methods compromise the “capsule” [10]. Formalin fixation causes the formation of chemical cross links of nucleic acids and proteins, thus resulting in degraded, fragmented DNA and RNA [11]. The current study sought to evaluate an alternative, non-crosslinking fixation reagent for prostate specimens. We demonstrate that fixation of prostates with PAXgene results in preservation of high quality DNA and RNA with preserved tissue architecture and staining characteristics.

Materials and methods

Tissue preparation

Four random prostates devoid of PCa were obtained at autopsy and prepared by the Human Tissue Resource Center (HTRC) at the University of Chicago. Before fixation, a punch biopsy was immediately embedded in OCT compound and frozen. The remainder of each specimen was bisected. Half was fixed in 10% buffered formalin for 24 hours using standard fixation techniques and embedded in paraffin. The other half was placed in PAXgene Tissue FIX, 150 mL, or adjusted to ~4 mL per gram of tissue, for either 24 or 48 hours (**Figure 1**). Specimens were then incubated in ethanol-diluted PAXgene Tissue STABILIZER, 150 mL, or adjusted to ~4 mL per gram of tissue, for at least three hours. PAXgene-fixed specimens

were embedded in paraffin using standard ethanol-based techniques, avoiding all exposure to formalin.

Similarly, as part of an Internal Review Board-approved protocol, specimens with PCa from four patients undergoing radical prostatectomy were fixed in PAXgene Tissue FIX for 48 hours and in PAXgene Tissue STABILIZER for at least three hours. They were then embedded in paraffin using standard ethanol-based techniques.

Hematoxylin and eosin (H&E) staining and immunohistochemistry (IHC)

Five μ m sections from PAXgene-fixed paraffin embedded (PAXPE) and formalin-fixed paraffin embedded (FFPE) sample blocks were stained with hematoxylin and eosin according to standard procedures. Two pathologists with expertise in genitourinary malignancies (T.A., G.P.) performed independent morphology assessment including overall morphology and nuclear, cytoplasmic and cell membrane details of the prostatic gland and stroma, while blinded to the fixation method.

For IHC, tissue sections were treated and stained using the automated VentanaBenchMark XT System (PSA) or the Leica Bond III System (P63 and P504). For PSA detection, anti-PSA antibody (DAKO, cat#A0562, Rabbit IgG, polyclonal, 1:1500) was applied on tissue sections and the antigen-antibody binding was detected with Ultraview detection kit from Ventana. For P63 detection, tissue sections were treated with Bond Epitope Retrieval Solution 2 (AR9640, Leica, Biosystems, EDTA, pH 9) for 20 minutes. Anti-P63 antibody (Biocaremedical, cat#CM163, clone: 4A4, mouse IgG2a, 1:100) was applied on tissue sections for 25 minutes. For P504 detection, tissue sections were treated with Bond Epitope Retrieval Solution 1 (AR9961, Leica, Biosystems, citrate buffer, pH 6) for 30 minutes. Anti-P504S antibody (Biocaremedical, cat#ACA-200BK, Rabbit IgG, polyclonal, 1:50) was applied on tissue sections for 25 minutes. For P63 and P504 the antigen-antibody binding was detected using Leica Bond Refine polymer detection system (Leica Biosystems, DS9800). Tissue sections were briefly immersed in hematoxylin for counterstaining. Images were

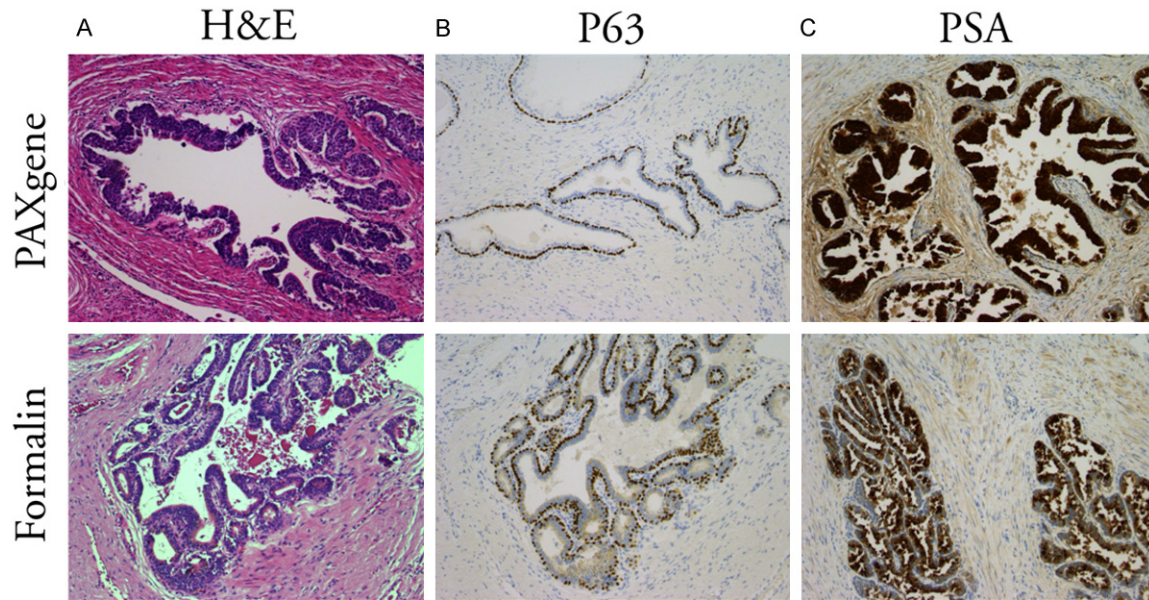


Figure 2. Preservation of morphology and antigenicity depending on fixation method. H&E staining (left), and P63 (center) and PSA (right) immunohistochemistry of prostate tissues from autopsy specimens fixed in PAXgene (top) for 48 hours or in formalin (bottom). Magnification X10.

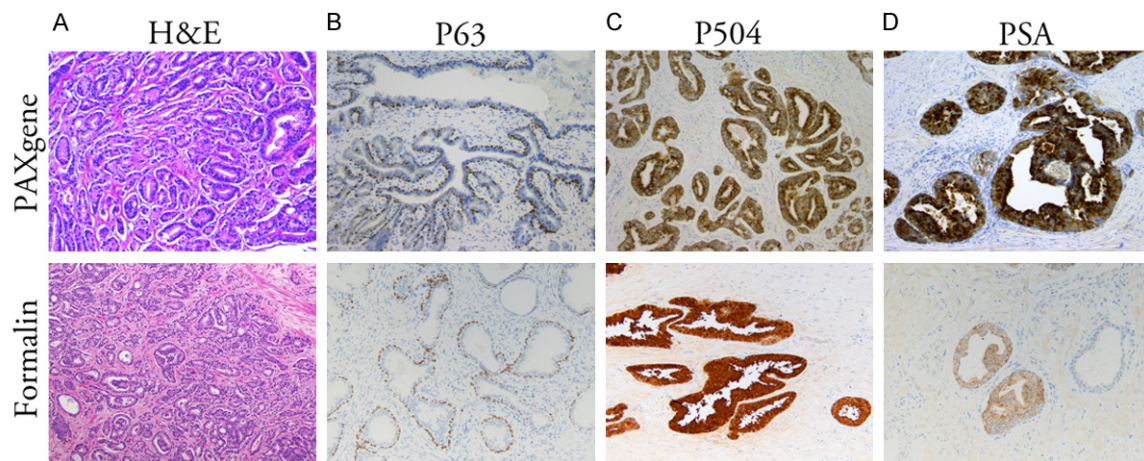


Figure 3. Prostate cancer staining. H&E staining (A) and P63 (B), P504S (C) and PSA (D) immunohistochemical staining of prostatectomy tissues fixed in PAXgene for 48 hours (top) or formalin (bottom). Magnification X10.

obtained on a DM6000B microscope (Leica) at 10x magnification.

Nucleic acids extraction

Genomic DNA and total RNA from PAXgene-fixed samples were isolated using PAXgene Tissue DNA kit (PreAnalytiX) and PAXgene Tissue miRNA kit (PreAnalytiX), respectively. Genomic DNA and total RNA from formalin-fixed samples were isolated using QIAamp DNA FFPE Tissue kit (Qiagen) and miRNeasy FFPE kit

(Qiagen), respectively. Genomic DNA and total RNA from frozen OCT embedded samples were isolated using DNeasy Blood and Tissue kit (Qiagen) and miRNeasy Mini kit (Qiagen), respectively. All extractions were performed using the manufacturer's recommended protocols. Extracted DNA concentration was measured by Qbit® fluorometric quantification (Life Technologies). DNA fragmentation and RNA integrity number (RIN) were evaluated by Screentape® electrophoresis system (Agilent

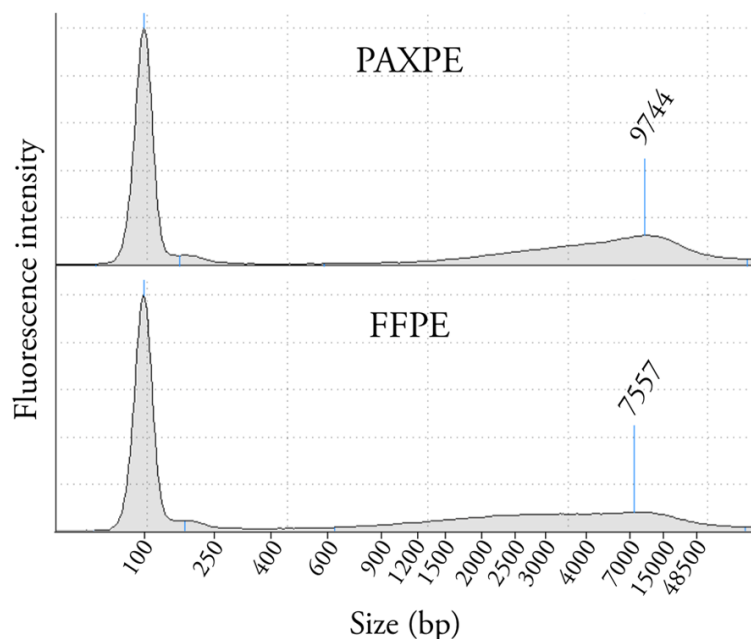


Figure 4. DNA fragment length depending on fixation method. Representative results of DNA fragments length between prostate tissue samples fixed in PAXgene for 48 hours or in formalin evaluated by Screentape® electrophoresis system. The electrophoregram for the FFPE sample shows smaller more degraded DNA as compared to the PAXPE sample.

Technologies). After extraction, DNA was stored at 4°C and RNA at -80°C.

Reverse transcription assay

For each sample, 1 µg of RNA was transcribed into cDNA using SuperScript VILO MasterMix (Life Technologies) according to the manufacturer's instructions. After reverse transcription, cDNA was stored at -20°C.

Real time PCR assay

For testing DNA integrity, one forward primer and six reverse primers ([Supplementary Table 1](#)) were designed to amplify six different transcript sequences with lengths of 88 to 817 base-pairs (bp) of the GAPDH gene. For testing cDNA integrity, one forward primer and five reverse primers ([Supplementary Table 1](#)) were designed to amplify five different transcript sequences with lengths of 95 to 983 bp of the β-Actin (ACTB) gene. Real time quantitative PCR assays were performed with 10 ng of genomic DNA and one microliter of 1:4 cDNA dilution, using iTaQ Universal SYBR Green Supermix (Bio-rad) according to the manufac-

turer's instructions. Reactions were performed on a ViiA 7 Real-Time PCR System (Life Technologies). Thermal cycling conditions were as followed: an activation step of 15 min at 95°C and 40 cycles of 15 s at 94°C, 30 s at 62°C and 30 s at 70°C. PCR reactions were run in triplicate. The acceptance criteria for single reactions were that the cycle threshold (CT) must be below 40 and the standard deviation in CT between triplicate reactions must be below 0.15. Melting curves had to be free of extraneous peaks. The significance of the difference between PAXPE and FFPE samples was tested using a paired Student's T-test.

Results

Preservation of tissue morphology and immunohistochemistry characteristics

To evaluate the preservation of tissue morphology under PAXgene fixation, sections of formalin-fixed, paraffin-embedded (FFPE) and PAXgene-fixed, paraffin-embedded (PAXPE) specimens were stained with hematoxylin and eosin. On pathology assessment, the preservation of tissue morphology in specimens was similar regardless of fixation agent used (**Figure 2A**). It was noted that using equivalent staining protocols, PAXgene-fixed tissue exhibits stronger eosinophilic stain, which does not affect the nuclear detail. Some compromise in morphology was noted in the autopsy sections fixed using either technique, which was likely due to extended post-mortem time prior to fixation.

Analysis of prostate tissue relies on immunohistochemical analysis using antibodies to Prostate Specific Antigen (PSA), the basal cell marker P63, and P504, a marker supporting a diagnosis of malignancy. Using standard IHC workflows, sections of PAXPE and FFPE prostates devoid of cancer were evaluated for PSA and P63. Comparable staining characteristics and staining intensity were observed between PAXPE and FFPE specimen for PSA and P63, (**Figure 2B, 2C**). PSA staining was detected as a

Preservation of prostate tissue in fidelity

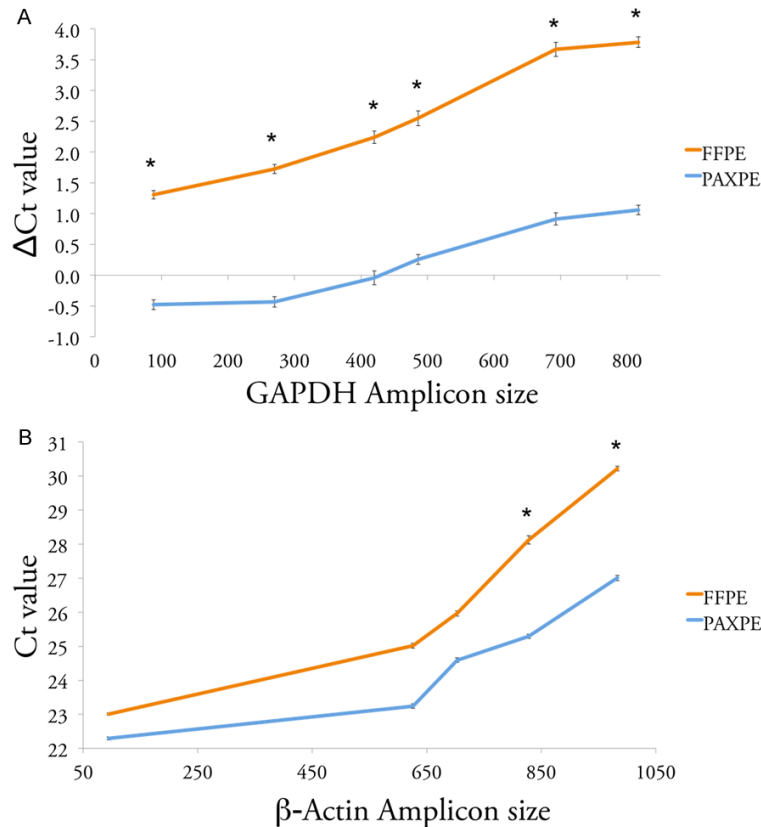


Figure 5. Preservation of DNA and RNA depending on fixation method. A. Preservation of DNA. Line plot showing the difference in cycle threshold (Ct) in real-time PCR targeting GAPDH gene DNA, between PAXgene-fixed and formalin-fixed prostate tissue samples, and the corresponding Frozen tissue. Each line represents the average values from 2 prostate specimens from autopsies fixed for 48 hours. *: statistical significance, P -value < 0.05. B. Preservation of RNA. Line plot showing threshold cycle (Ct) of real-time RT-PCR targeting β -Actin gene cDNA from PAXgene-fixed or formalin-fixed prostate tissue samples. Each line represents the average values from 2 prostate specimens from autopsies fixed for 48 hours. *: statistical significance, P -value < 0.05.

very intense stain in luminal cells of prostate glands whereas P63 showed as strong nuclear staining in the basal cells.

To confirm that there was minimal effect of PAXgene fixation on immunohistochemical analysis of tissue containing PCa, four prostatectomy specimens were fixed in PAXgene reagent for 48 hours. As seen with non-PCa tissue, H&E staining demonstrated that tissue morphology and architecture was well preserved (**Figure 3A**). PCa specimens underwent immunohistochemical evaluation for P504, in addition to PSA and P63. All three stained strongly, demonstrating high levels of immune

reactivity on PAXPE sections (**Figure 3B-D**).

Preservation of nucleic acid quality

To investigate the quality of nucleic acids from the different fixation methods, total DNA and RNA were extracted from each autopsy prostate specimen. Sample yield varied highly between specimens but did not show significant differences between the preservation methods (data not shown). For DNA from both types of tissue the average fragment length was 5Kb or more. PAXgene-fixed samples were ~25% longer than formalin-fixed ones (**Figure 4**). For RNA samples, RIN scores were below 2, and there was no significant difference between PAXgene- and Formalin-fixed tissues. Quality of the RNA extracted from frozen tissue sample from each autopsy specimen ranged from RIN 1.5 to 3.

The relative quality of DNA and RNA was evaluated by real-time PCR and reverse transcription real-time PCR, respectively. To measure the DNA template quality, as a

function of the fixation method, six primer pairs were designed to amplify fragments of increasing length of the GAPDH gene, from 88 bp to 817 bp. Consistently, DNA from PAXPE samples demonstrated 4-8 fold increased amplification efficiency (corresponding to 2-3 lower cycle threshold value) compared to FFPE samples (**Figure 5A**).

Five primer pairs were used to evaluate cDNA quality, producing amplicons of increasing length of the β -Actin house-keeping gene, from 95 bp to 983 bp. While RIN scores were low in both formalin and PAXgene tissues, there was a significant difference in cDNA template quality, with up to 8-fold increased amplification efficiency in PAXPE samples (**Figure 5B**).

Preservation of prostate tissue in fidelity

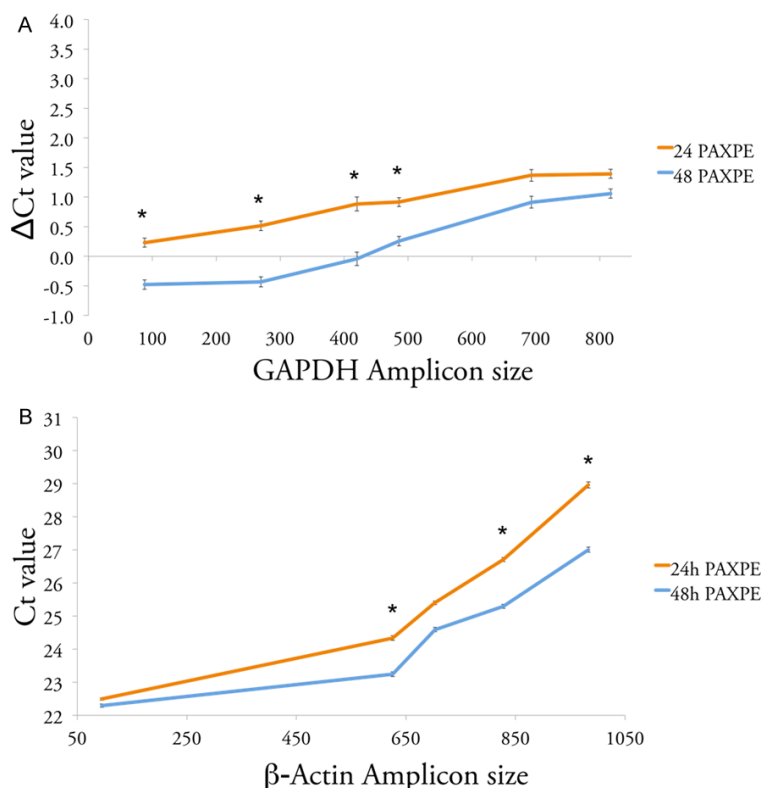


Figure 6. Preservation of DNA and RNA depending on PAXgene fixation time. A. Preservation of DNA. Line plot showing the difference in cycle threshold (Ct) in real-time PCR targeting GAPDH gene DNA, between prostate tissue samples fixed in PAXgene for 24 or 48 hours, and the corresponding Frozen tissue. Each line represents the average values from 2 prostate specimens from autopsies. *: statistical significance, P -value < 0.05 . B. Preservation of RNA. Line plot showing threshold cycle (Ct) of real-time RT-PCR targeting β -Actin gene cDNA from prostate tissue samples fixed in PAXgene for 24 or 48 hours. Each line represents the average values from 2 prostate specimens from autopsies. *: statistical significance, P -value < 0.05 .

Fixation with PAXgene for 48 hours requires alterations to the typical tissue processing workflow and a delay in availability of tissue for diagnostic studies. Therefore we also evaluated the quality of DNA and RNA from PAXPE prostate specimens fixed for 24 hours. As seen in **Figure 6**, quality of both DNA and RNA from these specimens was slightly inferior to that fixed for 48 hours.

Discussion

We report here that PAXgene fixation and stabilization reagents preserve morphology and immunoreactivity equivalent to formalin fixation, but with superior preservation of nucleic acid integrity. This technique can facilitate collection of tissue with high quality DNA and RNA, without compromising tissue architecture by

disrupting the capsule prior to fixation, and without limiting tissue collection to small samples obtained without prior full pathological review. This represents a particular advantage for studies evaluating the entire gland or correlating morphologic or staining features with molecular characteristics.

Hematoxylin and eosin stained PAXPE prostate tissues provided good representation of histologic features, similar to that of formalin. As previously described [12, 13] an EC FP7 project aimed to improve pre analytic procedures, the PAXgene Tissue System (PAXgene) both hematoxylin and eosin were modestly more intense in PAXPE samples than in the corresponding FFPE samples. Immunohistochemical analysis revealed that fixation with PAXgene did not alter immunoreactivity of antigens relevant for PCa as compared to formalin fixation. Staining and counterstaining were intense for all markers tested. Importantly, immunohistochemical analysis for this study did not require specific optimization of routine staining procedures for PAXPE

samples, in contrast to what has been previously reported [12].

Consistent with previous studies, we found that PAXgene reagents preserved both DNA and RNA quality significantly better than formalin [14-16]. Using real-time PCR on DNA and RNA obtained from prostate autopsy specimens, we showed that PAXPE samples have lower CT values compared to FFPE samples. This difference increased with amplicon length, indicating that PAXgene allows efficient amplification even of long fragments. Chromosomal rearrangements and fusion genes appear to play an important role in prostate cancer development and progression [4, 17-19]. Identification of these structural variants using next-generation sequencing is improved by sequencing across longer DNA inserts, thus increasing the physi-

cal coverage of the genome [20]. Sequencing libraries made from FFPE require short inserts to allow for amplification across cross-linked DNA. DNA from PAXPE tissue facilitates use of sequencing libraries more amenable to identifying the rearrangements characteristic of prostate cancer.

We demonstrate here significant differences in downstream assays using RNA extracted from PAXPE versus FFPE prostatectomy samples, though RNA from both tissue types were of low quality. It should be noted that RNA is likely to have substantially degraded during post-mortem time before fixation, which is supported by the low quality of RNA from matched frozen tissue.

A limitation of this study is that a portion of the study used autopsy specimens, which have already undergone some degree of post-mortem degradation prior to PAXgene and formalin fixation. This strategy was undertaken to avoid compromising the capsule of clinical samples prior to fixation. Prostatectomy specimens fixed wholly in PAXgene also demonstrated preserved morphology and staining characteristics. Although we demonstrate superior preservation of nucleic acids with PAXgene, the question of how long those features remain stable is still to be answered.

The prolonged duration of fixation (48 hours) may be problematic for the workflow of some centers. Though 24 hour fixation is inferior to 48 hour fixation in preserving DNA quality, it is superior to formalin and preserves staining characteristics and therefore is a suitable alternative. In addition, the routine use of PAXgene fixation increases financial costs. If prohibitively expensive for routine use, PAX gene fixation is an excellent method for preserving specimens of high research interest.

In this study we thus demonstrate the novel use of PAXgene for fixation of prostatectomy specimens. Its use improves the quality of molecular analysis without compromising histopathological analysis. PAXgene fixation may prove especially useful for assays for which longer amplicons are required, such as for chromosomal rearrangements, or for studies making use of more tissue than can be obtained with a simple punch.

Conclusions

Given the difficulty of identifying PCa on gross examination and the potential for compromise of the prostate “capsule” if tissue is divided prior to fixation, banking significant amounts of PCa tissue samples is challenging. The alternative, non-crosslinking fixation agent PAXgene preserves the quality of DNA and RNA isolated from fixed tissue without compromising the quality of tissue histology or staining characteristics. Incorporation of PAXgene fixation into banking protocols should facilitate the wider availability of high quality PCa tissues for next-generation sequencing and other studies.

Acknowledgements

This work was supported in part by the University of Chicago Cancer Center Support Grant (CA014599), and some PAXgene reagents were provided by Qiagen.

Disclosure of conflict of interest

None.

Address correspondence to: David J VanderWeele, Department of Medicine, University of Chicago, Chicago, IL USA. E-mail: djvanderweele@gmail.com

References

- [1] Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. *CA Cancer J Clin* 2014; 64: 9-29.
- [2] Mohler J, Bahnson RR, Boston B, Busby JE, D'Amico A, Eastham JA, Enke CA, George D, Horwitz EM, Huben RP, Kantoff P, Kawachi M, Kuettel M, Lange PH, Macvicar G, Plimack ER, Pow-Sang JM, Roach M 3rd, Rohren E, Roth BJ, Shrieve DC, Smith MR, Srinivas S, Twardowski P, Walsh PC. NCCN clinical practice guidelines in oncology: prostate cancer. *J Natl Compr Cancer Netw JNCCN* 2010; 8: 162-200.
- [3] Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS, Arora VK, Kaushik P, Cerami E, Reva B, Antipin Y, Mitsiades N, Landers T, Dolgalev I, Major JE, Wilson M, Socci ND, Lash AE, Heguy A, Eastham JA, Scher HI, Ruter VE, Scardino PT, Sander C, Sawyers CL, Gerald WL. Integrative genomic profiling of human prostate cancer. *Cancer Cell* 2010; 18: 11-22.
- [4] Berger MF, Lawrence MS, Demichelis F, Drier Y, Cibulskis K, Sivachenko AY, Sboner A, Esgueva R, Pflueger D, Sougnez C, Onofrio R, Carter SL, Park K, Habegger L, Ambrogio L, Fennell T, Par-

- kin M, Saksena G, Voet D, Ramos AH, Pugh TJ, Wilkinson J, Fisher S, Winckler W, Mahan S, Ardlie K, Baldwin J, Simons JW, Kitabayashi N, MacDonald TY, Kantoff PW, Chin L, Gabriel SB, Gerstein MB, Golub TR, Meyerson M, Tewari A, Lander ES, Getz G, Rubin MA, Garraway LA. The genomic complexity of primary human prostate cancer. *Nature* 2011; 470: 214-220.
- [5] Barbieri CE, Baca SC, Lawrence MS, Demichelis F, Blattner M, Theurillat JP, White TA, Stojanov P, Van Allen E, Stransky N, Nickerson E, Chae SS, Boysen G, Auclair D, Onofrio RC, Park K, Kitabayashi N, MacDonald TY, Sheikh K, Vuong T, Guiducci C, Cibulskis K, Sivachenko A, Carter SL, Saksena G, Voet D, Hussain WM, Ramos AH, Winckler W, Redman MC, Ardlie K, Tewari AK, Mosquera JM, Rupp N, Wild PJ, Moch H, Morrissey C, Nelson PS, Kantoff PW, Gabriel SB, Golub TR, Meyerson M, Lander ES, Getz G, Rubin MA, Garraway LA. Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat Genet* 2012; 44: 685-689.
- [6] Grasso CS, Wu YM, Robinson DR, Cao X, Dhanasekaran SM, Khan AP, Quist MJ, Jing X, Lonigro RJ, Brenner JC, Asangani IA, Ateeq B, Chun SY, Siddiqui J, Sam L, Anstett M, Mehra R, Prensner JR, Palanisamy N, Ryslik GA, Vandin F, Raphael BJ, Kunju LP, Rhodes DR, Pienta KJ, Chinnaiyan AM, Tomlins SA. The mutational landscape of lethal castration-resistant prostate cancer. *Nature* 2012; 487: 239-243.
- [7] Beltran H, Rickman DS, Park K, Chae SS, Sboner A, MacDonald TY, Wang Y, Sheikh KL, Terry S, Tagawa ST, Dhir R, Nelson JB, De La Taille A, Allory Y, Gerstein MB, Perner S, Pienta KJ, Chinnaiyan AM, Wang Y, Collins CC, Gleave ME, Demichelis F, Nanus DM, Rubin MA. Molecular Characterization of Neuroendocrine Prostate Cancer and Identification of New Drug Targets. *Cancer Discov* 2011; 1: 487-495.
- [8] Brenner JC, Ateeq B, Li Y, Yocum AK, Cao Q, Asangani IA, Patel S, Wang X, Liang H, Yu J, Palanisamy N, Siddiqui J, Yan W, Cao X, Mehra R, Sabolch A, Basur V, Lonigro RJ, Yang J, Tomlins SA, Maher CA, Elenitoba-Johnson KSJ, Hussain M, Navone NM, Pienta KJ, Varambally S, Feng FY, Chinnaiyan AM. Mechanistic rationale for inhibition of poly(ADP-ribose) polymerase in ETS gene fusion-positive prostate cancer. *Cancer Cell* 2011; 19: 664-678.
- [9] Samaratunga H, Montironi R, True L, Epstein JI, Griffiths DF, Humphrey PA, van der Kwast T, Wheeler TM, Srigley JR, Delahunt B, Egevad L. International Society of Urological Pathology (ISUP) Consensus Conference on Handling and Staging of Radical Prostatectomy Specimens. Working group 1: specimen handling. *Mod Pathol* 2011; 24: 615.
- [10] Esgueva R, Park K, Kim R, Kitabayashi N, Barbieri CE, Dorsey PJ Jr, Abraham C, Banerjee S, Leung RA, Tewari AK, Terry S, Shevchuk MM, Rickman DS, Rubin MA. Next-generation prostate cancer biobanking: toward a processing protocol amenable for the International Cancer Genome Consortium. *Diagn Mol Pathol Am J Surg Pathol Part B* 2012; 21: 61-68.
- [11] Van Beers EH, Joosse SA, Ligtenberg MJ, Fles R, Hogervorst FBL, Verhoef S, Nederlof PM. A multiplex PCR predictor for aCGH success of FFPE samples. *Br J Cancer* 2005; 94: 333-337.
- [12] Kap M, Smedts F, Oosterhuis W, Winther R, Christensen N, Reischauer B, Viertler C, Groelz D, Becker KF, Zatloukal K, Langer R, Slotta-Huspenina J, Bodo K, de Jong B, Oelmüller U, Riegman P. Histological Assessment of PAXgene Tissue Fixation and Stabilization Reagents. *PLoS One* 2011; 6: e27704.
- [13] Gündisch S, Slotta-Huspenina J, Verderio P, Ciniselli CM, Pizzamiglio S, Schott C, Drecoll E, Viertler C, Zatloukal K, Kap M, Riegman P, Esposito I, Specht K, Babaryka G, Asslaber M, Bodó K, Bakker M den, Hollander J den, Fend F, Neumann J, Reu S, Perren A, Langer R, Lugli A, Becker I, Richter T, Kayser G, May AM, Carneiro F, Lopes JM, Sobin L, Höfler H, Becker KF. Evaluation of colon cancer histomorphology: a comparison between formalin and PAXgene tissue fixation by an international ring trial. *Virchows Arch* 2014; 465: 509-519.
- [14] Ergin B, Meding S, Langer R, Kap M, Viertler C, Schott C, Ferch U, Riegman P, Zatloukal K, Walch A, Becker KF. Proteomic Analysis of PAXgene-Fixed Tissues. *J Proteome Res* 2010; 9: 5188-5196.
- [15] Belloni B, Lambertini C, Nuciforo P, Phillips J, Bruening E, Wong S, Dummer R. Will PAXgene substitute formalin? A morphological and molecular comparative study using a new fixative system. *J Clin Pathol* 2013; 66: 124-135.
- [16] Viertler C, Groelz D, Gündisch S, Kashofer K, Reischauer B, Riegman PHJ, Winther R, Wyrich R, Becker KF, Oelmüller U, Zatloukal K. A New Technology for Stabilization of Biomolecules in Tissues for Combined Histological and Molecular Analyses. *J Mol Diagn* 2012; 14: 458-466.
- [17] Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, Varambally S, Cao X, Tchinda J, Kuefer R, Lee C, Montie JE, Shah RB, Pienta KJ, Rubin MA, Chinnaiyan AM. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* 2005; 310: 644-648.
- [18] Baca SC, Prandi D, Lawrence MS, Mosquera JM, Romanell A, Drier Y, Park K, Kitabayashi N, MacDonald TY, Ghandi M, Van Allen E, Kryukov GV, Sboner A, Theurillat JP, Soong TD, Nicker-

Preservation of prostate tissue in fidelity

son E, Auclair D, Tewari A, Beltran H, Onofrio RC, Boysen G, Guiducci C, Barbieri CE, Cibulskis K, Sivachenko A, Carter SL, Saksena G, Voet D, Ramos AH, Winckler W, Cipicchio M, Ardlie K, Kantoff PW, Berger MF, Gabriel SB, Golub TR, Meyerson M, Lander ES, Elemento O, Getz G, Demichelis F, Rubin MA, Garraway LA. Punctuated evolution of prostate cancer genomes. *Cell* 2013; 153: 666-677.

- [19] Rubin MA, Maher CA, Chinnaiyan AM. Common Gene Rearrangements in Prostate Cancer. *J Clin Oncol* 2011; 29: 3659-3668.
- [20] Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. *Nat Rev Genet* 2010; 11: 685-696.

Preservation of prostate tissue in fidelity

Supplementary Table 1. Sequences of primers used for real time PCR assay

Name	Sequence (5'-3')
GAPDH-F	AGTCCCAGAAACAGGAGGT
GAPDH-R-88 bp	AGAGCGCGAAAGGAAAGAA
GAPDH-R-270 bp	GTTAGTCACCGGCAGGCTTT
GAPDH-R-420 bp	CACAAGAGGACCTCCATAAACC
GAPDH-R-486 bp	ACCCATGACTCAGCTTCTCC
GAPDH-R-693 bp	GTTTCCGGAAGACGGAATG
GAPDH-R-817 bp	GGGAGCACAGGTAAGTGCAT
B-ACTIN-F	AAATCTGGCACCACACCTTC
B-ACTIN-R-95 bp	GTTGGCCTTGGGGTTTCAG
B-ACTIN-R-626 bp	AGAGGCGTACAGGGATAGCA
B-ACTIN-R-703 bp	AGGGCATACCCTCGTAGAT
B-ACTIN-R-828 bp	TAATGTCACGCACGATTTC
B-ACTIN-R-983 bp	AGGCTGGAAGAGTGCCTCAG

RESEARCH

Open Access



Circulating tumor cells capture disease evolution in advanced prostate cancer

Justin Lack¹, Marc Gillard², Maggie Cam¹, Gladell P. Paner³ and David J. VanderWeele^{4,5*} 

Abstract

Background: Genetic analysis of advanced cancer is limited by availability of representative tissue. Biopsies of prostate cancer metastasized to bone are invasive with low quantity of tumor tissue. The prostate cancer genome is dynamic, however, with temporal heterogeneity requiring repeated evaluation as the disease evolves. Circulating tumor cells (CTCs) offer an alternative, “liquid biopsy”, though single CTC sequencing efforts are laborious with high failure rates.

Methods: We performed exome sequencing of matched treatment-naïve tumor tissue, castrate resistant tumor tissue, and pooled CTC samples, and compared mutations identified in each.

Results: Thirty-seven percent of CTC mutations were private to CTCs, one mutation was shared with treatment-naïve disease alone, and 62% of mutations were shared with castrate-resistant disease, either alone or with treatment-naïve disease. An acquired nonsense mutation in the *Retinoblastoma* gene, which is associated with progression to small cell cancer, was identified in castrate resistant and CTC samples, but not treatment-naïve disease. This timecourse correlated with the tumor acquiring neuroendocrine features and a change to neuroendocrine-specific therapy.

Conclusions: These data support the use of pooled CTCs to facilitate the genetic analysis of late stage prostate cancer.

Keywords: Circulating tumor cells, Castrate resistant prostate cancer, Tumor evolution, Neuroendocrine prostate cancer

Background

Prostate cancer biology is initially dominated by activity of the androgen receptor (AR), and androgen deprivation therapy (ADT) is the backbone of therapy for those with metastatic prostate cancer. Castrate resistance is nearly universal, however, requiring additional management decisions. The heterogeneity of primary prostate cancer is well established [1–3], and polyclonal seeding of metastatic sites, and metastasis-to-metastasis spread, appear to be possible, if not common [4]. Moreover, the prostate cancer genome is dynamic, with different clones dominating over time in response to different lines of therapy [5, 6]. As the number of approved agents that

prolong survival increases, the dynamic clonal nature of this disease is likely to become a greater issue. For example, there appears to be a new entity, intermediate between conventional adenocarcinoma and classic neuroendocrine disease, that arises with resistance to newer AR-targeted agents [7].

Circulating tumor cells (CTCs) and cell free DNA (cfDNA) are emerging as non-invasive alternative means to interrogate the genetics of late stage disease rather than invasive and/or risky biopsy of metastatic disease. They are complimentary methods, each with advantages and disadvantages. The dilute tumor fraction of cfDNA in most patients requires very high sequencing coverage to detect mutations, typically limiting analysis to a small number of targeted regions [5]. On the other hand, amplifying DNA from single circulating tumor cells is inconsistent and low yield. Successful efforts from individual cases have sequenced four of 99 collected CTCs,

*Correspondence: david.vanderweele@nih.gov

⁴ Laboratory for Genitourinary Pathogenesis, Center for Cancer Research, National Cancer Institute, 37 Convent Drive, Rm 1066A, Bethesda, MD 20892, USA

Full list of author information is available at the end of the article

or pooled the results from 19 separately sequenced CTCs [8, 9]. We present here whole exome sequencing of pooled CTCs, with matched treatment-naïve and castrate resistant tissue, demonstrating identification in pooled CTCs of clinically relevant mutations acquired late in the course of disease.

Methods

Patient recruitment and patient samples

The patient provided written informed consent and was enrolled in the “Prostate Cancer Sample Collection” protocol of the University of Chicago Medical Center, which was approved by the University of Chicago ethics committee (approval reference number 13-1295). Biopsy specimens were processed per clinical protocols including hematoxylin and eosin stain and immunohistochemical analysis of PSA expression. Circulating tumor cells were collected as previously described [10]. In short, mononuclear cells were enriched from 15 ml peripheral blood and stained with an Alexa-488 conjugated EpCAM antibody (Biolegend, 1:100) and a QDot800 conjugated CD45 antibody (Invitrogen, 1:100). CTCs were isolated by FACS-sorting EpCAM+/CD45− cells on a MoFlo XDP flow-sorting machine. One thousand WBCs were isolated by FACS-sorting EpCAM−/CD45+ cells as representative of germline DNA.

Whole genome amplification and exome sequencing

Isolated EpCAM+/CD45− CTCs were divided into three equal samples of 500 cells each and independently subjected to whole genome amplification (WGA) through multidisplacement amplification using REPLI-G (Qiagen), following the manufacturer’s instructions, as were EpCAM−/CD45+ WBCs representative of germline. The quality of amplification was evaluated through PCR amplification of eight targets [11] to evaluate uniform amplification across multiple chromosomes, and length of amplification product. Primers used are listed in Additional file 1. Exome sequencing libraries were prepared using NEB Next Ultra (New England Biolabs) kit following the manufacturer’s instructions. Sequencing was performed on a HiSeq 2000 (Illumina).

Sequencing analysis

Illumina sequencing data was mapped to the GRCh37 version of the human reference genome using BWA-MEM [12] and further processed following the GATK Best Practices [13]. Somatic variants were called for each tumor or CTC sample in a paired manner using MuTect [14] in high confidence mode. Potential germline variants were removed by excluding positions with germline coverage <20×, variant reads in the germline, or variants present in dbSNP, and remaining variants were enriched

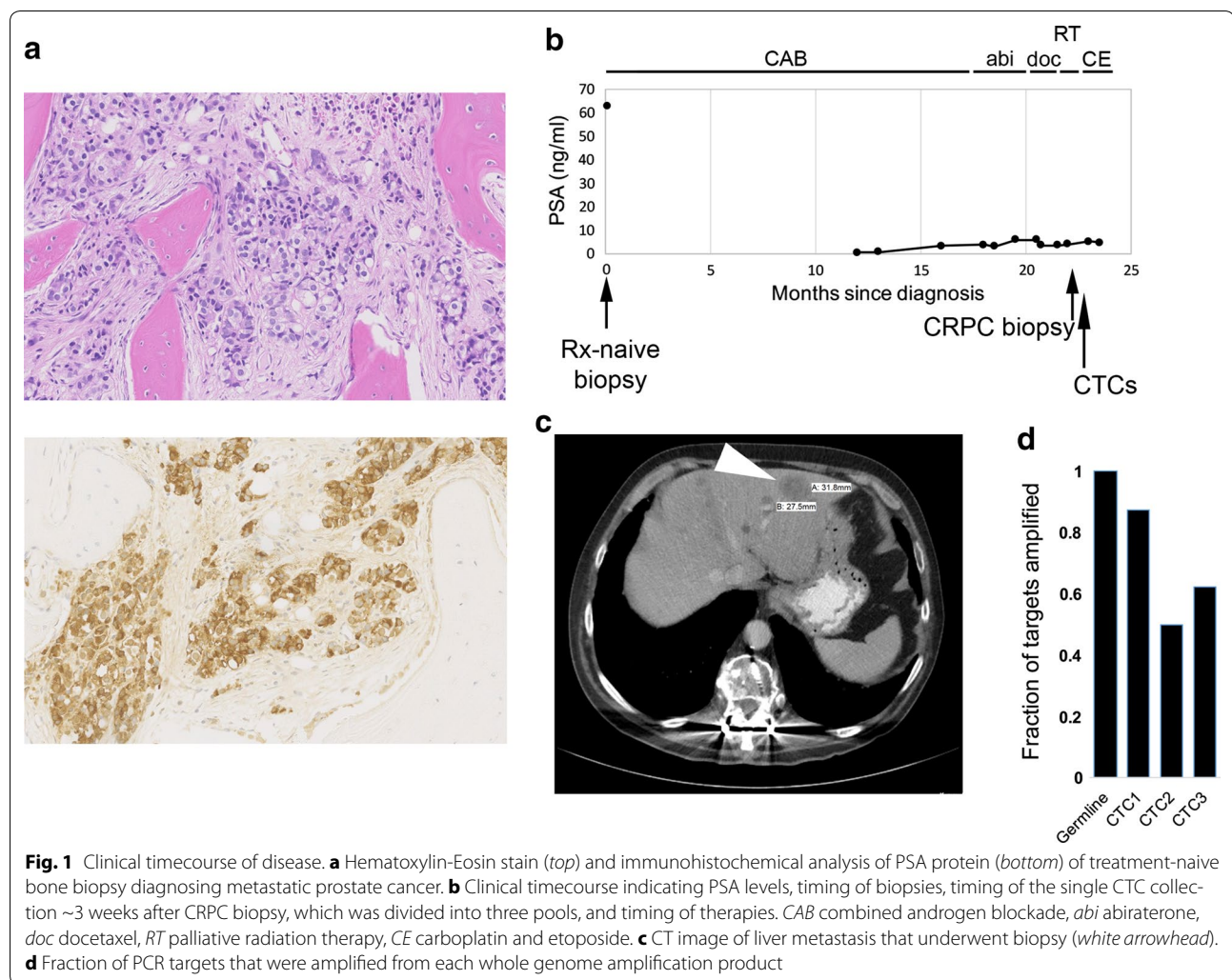
for highest confidence by including only variants in the exome and with 5 or more supporting reads comprising 10% or more of all reads. Somatic mutations passing all filters (listed in Additional file 2) in a given sample were then examined for the presence/absence of supporting reads in other samples using samtools mpileup [15, 16]. Mutations were visually verified by examining supporting reads using Alview [17], and all somatic mutations were annotated using AVIA [18].

Results

A 70 year old patient was diagnosed with metastatic prostate cancer, confirmed with biopsy of a bone metastasis (Fig. 1a). He underwent androgen deprivation therapy but eventually developed castrate resistant prostate cancer (CRPC). Additional conventional therapies yielded short-lived responses and subsequent progression of disease (Fig. 1b). A liver biopsy performed 22 months after the initial diagnosis confirmed metastatic prostate cancer (Fig. 1c), but with neuroendocrine and small cell features, which was not appreciated on the initial biopsy. Given the change in dominant histology he started treatment with carboplatin and etoposide, but after two cycles he entered hospice care.

To evaluate the extent to which the current disease biology can be discovered through CTCs, CTCs were collected from 15 ml of peripheral blood within a month of the liver biopsy. CTCs were enriched by FACS-sorting EpCAM+/CD45− cells as previously described [10] and divided into three equal fractions. Multidisplacement amplification was used to amplify the entire genome of each fraction independently. To evaluate for loss of coverage due to lack of amplification and for length of amplified fragments, the amplified CTC genomes were evaluated with endpoint PCR for eight targets of varying lengths across six different chromosomes (Fig. 1d). More targets were able to be amplified from one CTC pool (CTC1) than the other two pools (CTC2, CTC3). All three samples were used for exome sequencing.

Whole exome sequencing was performed on a germline sample and five tumor samples: initial diagnostic, treatment-naïve biopsy of an ischial metastasis; liver biopsy of CRPC; and three samples of pooled CTCs divided from a single CTC collection. The fractions of PCR duplicates were all 0.25 or less (Fig. 2a). Adequate coverage was obtained for the germline sample, the tissue samples, and CTC1, samples that were predicted to perform well by initial evaluation, with coverage of over 98% of the genome, median depths of 44–106×, and 10× coverage of 90% or more of the exome (Fig. 2b). The lower quality CTC samples covered just 37 and 40% of the exome at 10×. Due to their poor quality they were left out of downstream analysis. Mutation patterns were



consistent in the treatment-naïve, CRPC, and CTC sample, including nucleotide substitutions (Fig. 2c) and type of mutation (Fig. 2d).

The allele frequency was comparable for mutations in the tissue specimens, but higher in the CTCs (Fig. 3a). Moreover, while trunk mutations (those found in both treatment-naïve and CRPC tissue samples) were near 50% allele frequency, suggesting close to 100% pure CTCs all with one allele mutated, branch (shared with one tissue sample) and leaf (found only in CTCs) mutations were at lower allele frequency, indicating genetic heterogeneity among CTCs (Fig. 3b).

There was considerable genetic heterogeneity among tumor tissue samples, with just 8% of mutations being trunk mutations, and 25% being found in two samples. Thirty-six percent of mutations from CRPC were not initially identified in the treatment-naïve sample but had at least two reads supporting their presence, indicating

enrichment in the CRPC tissue of subclonal populations from the treatment-naïve tissue. The reverse—mutations identified in the treatment-naïve sample with low frequency supporting reads in CRPC—was not identified (Fig. 3c, d). The pooled CTCs identified 71% of mutations shared by treatment-naïve and CRPC tissue samples, suggesting most mutations at high allele frequency (early mutations) were represented in the CTCs (Fig. 3e). Examining the two low quality CTC pools, each pool identified 22% of trunk mutations, consistent with their 10x coverage being 37–40% of the genome. All the trunk mutations identified in the low quality pools were also identified in the high quality pool, suggesting consistency in the mutations able to be identified by CTCs.

The CTCs had the highest fraction of mutations shared with at least one other sample, at 64% (Fig. 3f). Fifty percent were trunk mutations, shared with both tissue samples, and an additional 14% were shared with the CRPC

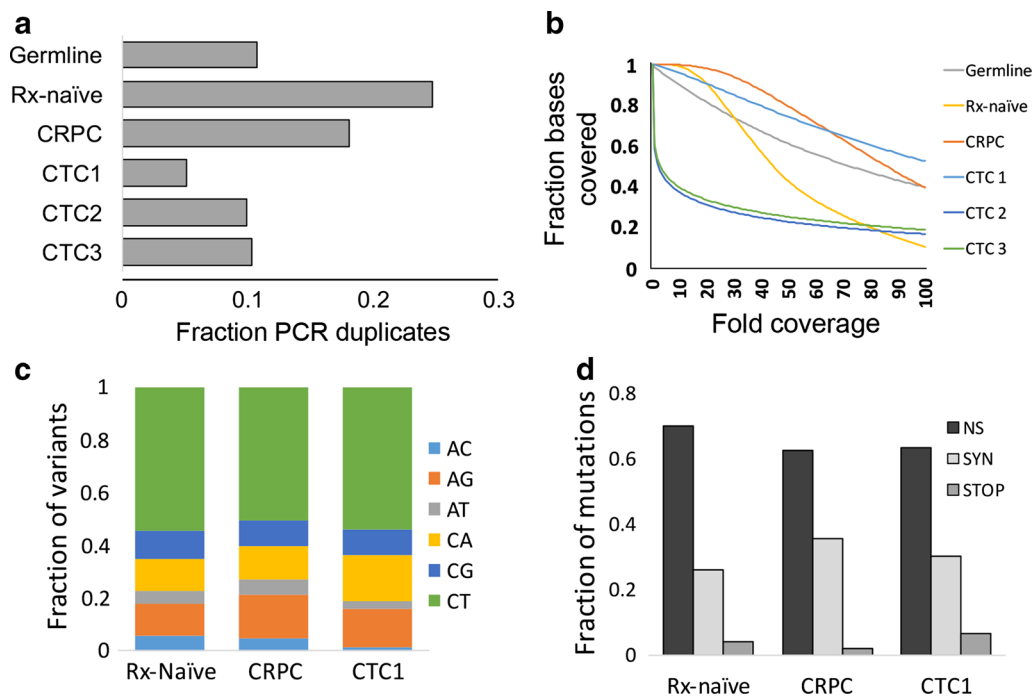


Fig. 2 Similar sequencing library characteristics from tissue and one of three CTC pool sequencing libraries. **a** Fraction of sequencing libraries consisting of PCR duplicates. **b** Coverage plots for germline, treatment-naïve (Rx-naïve) tissue, advanced cancer (CRPC) tissue, and three CTC pool sequencing libraries. **c** Transitions and transversions and **d** mutation types among the variants identified in the two tissue and CTC1 sequencing library

sample alone. A single mutation was shared between CTCs and treatment-naïve tissue and not found in CRPC tissue. Though the other two CTC samples had poor exome coverage and were not of sufficient quality to identify mutations a priori, 33% of mutations identified in the high quality CTC sample were also identified in one of the low quality CTC samples.

One of the mutations shared between the CRPC tissue sample and CTCs was a premature stop codon in the Retinoblastoma (RB1) gene (Fig. 4a), which is associated with isolated bilateral retinoblastoma and meningioma [19]. The appearance of this RB1 mutation coincided with loss of adenocarcinoma features, including PSA expression (compare Fig. 4b, top and middle panels, to Fig. 1a), and gain of neuroendocrine features in the tumor, which was confirmed by synaptophysin expression (Fig. 4b, bottom panel). This phenotypic change toward therapy-emergent neuroendocrine prostate cancer prompted a change in management strategy away from standard of care for CRPC.

Discussion

The prostate cancer genome is heterogeneous, both between and within the multiple foci characteristic of primary disease. The clonal architecture of advanced

disease is dynamic, with new clones gaining dominance in response to new therapies. Combined, these necessitate repeated genomic or molecular assessments over the course of disease in order to have a complete, current understanding of a patient's personalized disease.

Circulating tumor cells offer a non-invasive mechanism to repeatedly evaluate the shifting dynamics of disease. This has demonstrated clinically meaningful evaluations of specific alterations [20, 21]. However, there have been few direct correlations between treatment-naïve tissue, CTCs, and contemporaneous CRPC tissue, and wider scale genomic evaluations are still in early stages. Lohr et al. used whole genome sequencing to evaluate quality of amplified DNA followed by exome sequencing of 19 individual CTCs to demonstrate late divergence of CTCs and a previously resected lymph node [8]. Jiang et al. used laser capture microdissection to capture and evaluate 99 individual CTCs collected over five blood collections, identifying four individual CTCs with high quality DNA and eight with moderate quality [9]. The CTCs in that study identified 15% of trunk SNVs, with supporting reads for an additional 14% of reads.

Using our pooled CTC strategy, we generated successful sequencing libraries from 33% of samples. The

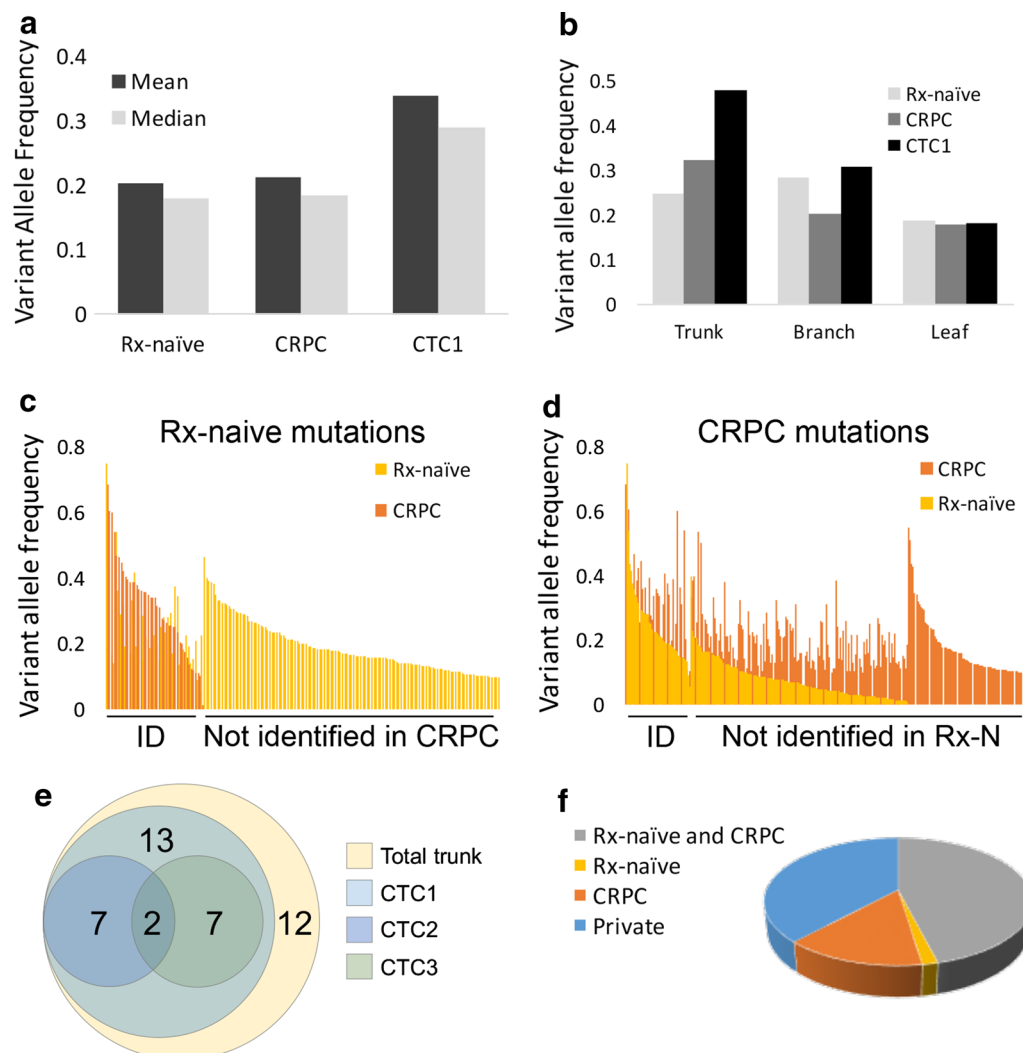
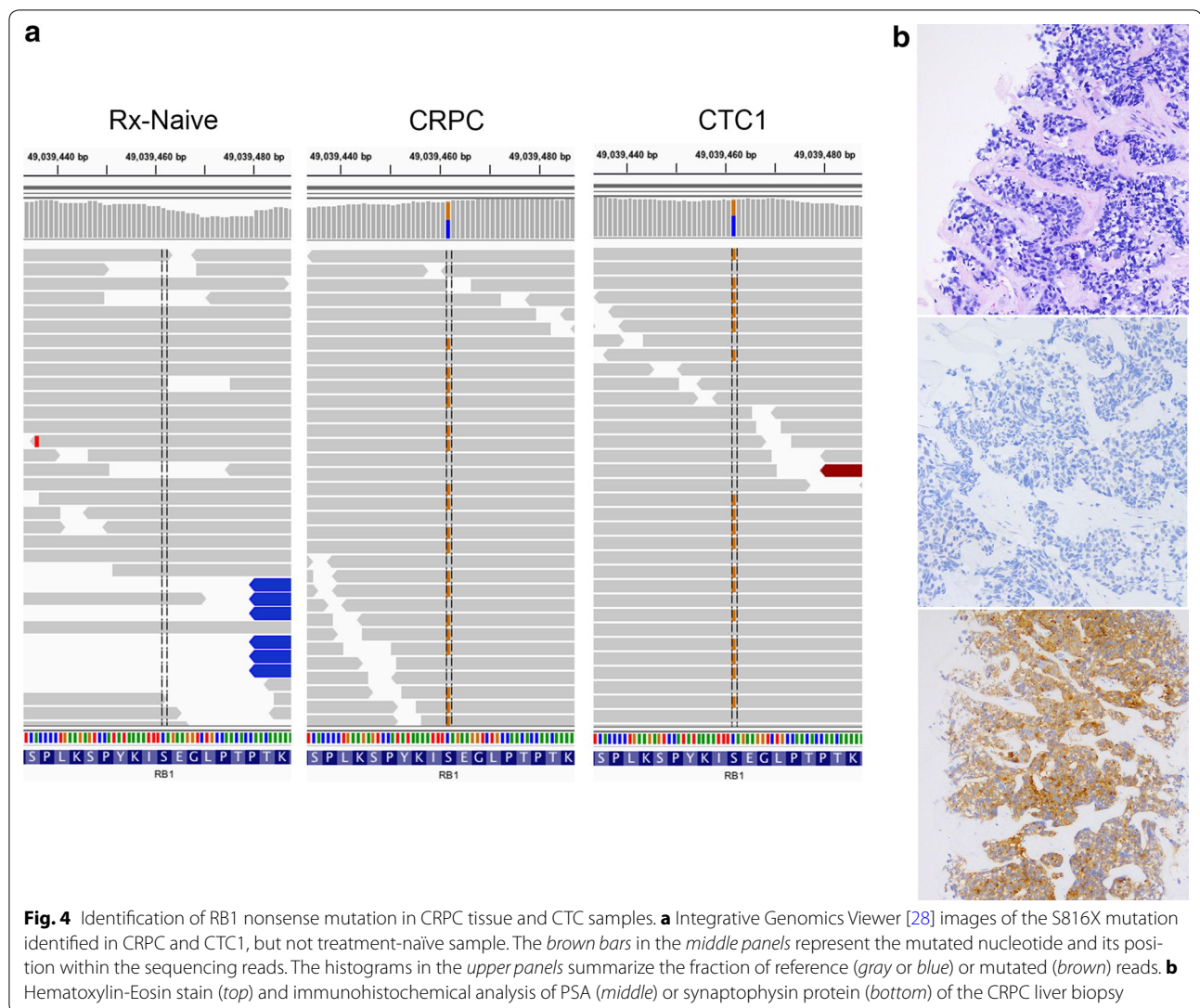


Fig. 3 Circulating tumor cells pool captures majority of trunk mutations and additional mutations from metastatic disease. **a** Variant allele frequencies for treatment-naïve, CRPC, and CTC1 mutations. **b** Average variant allele frequencies of trunk, branch, and leaf mutations in two tissue samples and one CTC sample. **c** Allele fraction in CRPC tissue of mutations identified in treatment-naïve tissue, and **d** vice versa. On the left are mutations identified in both samples. On the right are mutations identified in only sample. In **d**, many of the mutations not identified independently in treatment-naïve tissue had evidence they were present at low allele frequency. **e** Number of trunk mutations identified in three pooled CTC sequencing libraries. **f** Fraction of CTC mutations shared with other tissue samples

sequencing demonstrated high correlation with tissue samples, confirming the biological relevance of the CTC exomes, identifying 71% of trunk mutations, along with additional mutations acquired later in disease. This includes a clinically meaningful mutation in RB1 which likely contributed to a change in phenotype to neuroendocrine features, prompting a change in management strategies. The RB1 gene is altered in nearly 9% of advanced prostate cancer cases, through deletion, frameshift mutations, and introductions of premature stop codons [22]. Beltran et al. compared advanced prostate neuroendocrine and adenocarcinoma,

demonstrating that RB1 alterations are significantly enriched in advanced prostate cancer with neuroendocrine features (70% altered) compared to that with pure adenocarcinoma features (32% altered) [23]. Loss of RB1 function is common in primary small cell cancer of the prostate or lung, and in animal models it promotes development of small cell carcinoma [24, 25].

The clonal relationship among all three specimens suggest that neuroendocrine disease arose from adenocarcinoma, rather than being a coincident, independent clone. In addition, the high frequency of mutations in CRPC tissue that were present at low frequency in treatment-naïve



tissue supports the idea that advanced disease, including neuroendocrine disease, arises from subclonal population(s) in the initial specimen. Of note, though CTCs from patients with neuroendocrine prostate cancer are more frequently nonclassical than those with patients with adenocarcinoma (17), the RB1 mutation was identified in classical EpCAM+ CTCs.

Fewer mutations were identified in CTCs than in treatment-naïve or CRPC tissue samples. The significance of this is unclear. It may be that the limited number of CTCs was unable to capture the extensive diversity of clones comprising disease in the tissue. Alternatively, it may be that CTCs represent a limited number of aggressive and clinically relevant clones. The CTCs were not clonal, as evidenced by the presence of branch and leaf mutations. This genetic heterogeneity among CTCs is supported by

Massard et al. [26] based on a single genomic alteration, the ERG alteration pattern. The extent to which CTCs represent all the relevant subclones needs to be explored further.

There are several advantages of a pooled CTC strategy over single CTC sequencing, including availability of resources. Our strategy relied on FACS-sorting and whole genome amplification using a commercially available kit, which are readily available to most researchers. We did not require laser capture microdissection or robotic micromanipulation. Disadvantages include applicability limited to patients with a higher burden of CTCs and inability to fully characterize heterogeneity at the cellular level.

While we had a much higher success rate sequencing pooled CTCs than has been reported with single CTC

sequencing, only one of three pools provided high quality data. This may have been due in part to our multiple displacement amplification strategy [27]. This was chosen for the low error rate of its polymerase, but it may be less reliable in amplifying the majority of the genome compared to PCR-based methods.

As the number of effective therapies used for treatment of advanced prostate cancer increases, there is an increasing appreciation of the dynamic nature of the genomics of advanced disease in response to therapeutic pressure. We demonstrate here that sequencing pooled CTCs is a feasible, noninvasive, and informative way to evaluate the current molecular features of advanced disease.

Conclusions

The histology, behavior, and genomics of advanced prostate cancer evolve in response to therapeutic pressure. Pooled CTCs are a feasible, non-invasive way to interrogate the molecular characteristics of advanced prostate cancer.

Additional files

Additional file 1. PCR primer sequences.

Additional file 2. High confidence somatic mutations.

Authors' contributions

JL planned and performed analysis and helped write the manuscript; MG planned and performed experiments; MC planned analysis; GPP planned the experiments and reviewed pathology; DVW planned experiments, performed analysis, and wrote the manuscript. All authors read and approved the final manuscript.

Author details

¹ Center for Cancer Research Collaborative Bioinformatics Resource, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20892, USA. ² Department of Surgery, University of Chicago, Chicago, IL 60615, USA. ³ Department of Pathology, University of Chicago, Chicago, IL 60615, USA. ⁴ Laboratory for Genitourinary Pathogenesis, Center for Cancer Research, National Cancer Institute, 37 Convent Drive, Rm 1066A, Bethesda, MD 20892, USA. ⁵ Department of Medicine, University of Chicago, Chicago, IL 60615, USA.

Acknowledgements

We would like to thank the patient for his participation in this study. We would also like to thank Andrea Pontier for sequencing library preparation, Alexa Campanile and Julie Gruszczynski for clinical data management, and the University of Chicago Functional Genomics Facility for exome sequencing support.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and its additional files.

Consent for publication

Implicit within study enrollment.

Ethics approval and consent to participate

The patient provided written informed consent and was enrolled in the "Prostate Cancer Sample Collection" protocol of the University of Chicago Medical Center, which was approved by the University of Chicago ethics committee (approval reference number 13-1295).

Funding

This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs, through the Prostate Cancer Research Program under Award No. W81XWH-13-1-0451 (DVW). Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the Department of Defense. The U.S. Army Medical Research Acquisition Activity, 820 Chandler Street, Fort Detrick MD 21702-5014 is the awarding and administering acquisition office. This work was also supported by the University of Chicago Cancer Center Support Grant P30 CA014599 (DVW), and the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research (DVW). This work utilized the computational resources of the NIH HPC Biowulf cluster (<http://hpc.nih.gov>).

Received: 19 October 2016 Accepted: 3 February 2017

Published online: 23 February 2017

References

- VanderWeele DJ, Brown CD, Taxy JB, Gillard M, Hatcher DM, Tom WR, et al. Low-grade prostate cancer diverges early from high grade and metastatic disease. *Cancer Sci*. 2014;105:1079–85.
- Boutros PC, Fraser M, Harding NJ, de Borja R, Trudel D, Lalonde E, et al. Spatial genomic heterogeneity within localized, multifocal prostate cancer. *Nat Genet*. 2015;47:736–45.
- Cooper CS, Eeles R, Wedge DC, Van Loo P, Gundem G, Alexandrov LB, et al. Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and morphologically normal prostate tissue. *Nat Genet*. 2015;47:367–72.
- Gundem G, Van Loo P, Kremeyer B, Alexandrov LB, Tubio JMC, Papaemmanuil E, et al. The evolutionary history of lethal metastatic prostate cancer. *Nature*. 2015;520:353–7.
- Carreira S, Romanel A, Goodall J, Grist E, Ferraldeschi R, Miranda S, et al. Tumor clone dynamics in lethal prostate cancer. *Sci Transl Med*. 2014;6:254ra125.
- Nakazawa M, Lu C, Chen Y, Paller CJ, Carducci MA, Eisenberger MA, et al. Serial blood-based analysis of AR-V7 in men with advanced prostate cancer. *Ann Oncol Off J Eur Soc Med Oncol ESMO*. 2015;26:1859–65.
- Small EJ, Huang J, Youngren J, Sokolov A, Aggarwal RR, Thomas G, et al. Characterization of neuroendocrine prostate cancer (NEPC) in patients with metastatic castration resistant prostate cancer (mCRPC) resistant to abiraterone (Abi) or enzalutamide (Enz): preliminary results from the SU2C/PCF/AACR West Coast Prostate Cancer Dream Team (WCDT). *ASCO Meet Abstr*. 2015;33:5003.
- Lohr JG, Adalsteinsson VA, Cibulskis K, Choudhury AD, Rosenberg M, Cruz-Gordillo P, et al. Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. *Nat Biotechnol*. 2014;32:479–84.
- Jiang R, Lu Y-T, Ho H, Li B, Chen J-F, Lin M, et al. A comparison of isolated circulating tumor cells and tissue biopsies using whole-genome sequencing in prostate cancer. *Oncotarget*. 2015;6:44781–93.
- Reyes EE, VanderWeele DJ, Isikbay M, Duggan R, Campanile A, Stadler WM, et al. Quantitative characterization of androgen receptor protein expression and cellular localization in circulating tumor cells from patients with metastatic castration-resistant prostate cancer. *J Transl Med*. 2014;12:313.
- van Beers EH, Joosse SA, Ligtenberg MJ, Fles R, Hogervorst FBL, Verhoef S, et al. A multiplex PCR predictor for aCGH success of FFPE samples. *Br J Cancer*. 2006;94:333–7.
- Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv13033997 Q-Bio* [Internet]. 2013. <http://arxiv.org/abs/1303.3997>. Accessed 25 May 2016.

13. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20:1297–303.
14. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol.* 2013;31:213–9.
15. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinforma Oxf Engl.* 2009;25:2078–9.
16. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinforma Oxf Engl.* 2011;27:2987–93.
17. Finney RP, Chen Q-R, Nguyen CV, Hsu CH, Yan C, Hu Y, et al. Alview: portable software for viewing sequence reads in BAM formatted files. *Cancer Inform.* 2015;14:105–7.
18. Vuong H, Che A, Ravichandran S, Luke BT, Collins JR, Mudunuri US. AVIA v2.0: annotation, visualization and impact analysis of genomic variants and genes. *Bioinformatics.* 2015;31:2748–50.
19. Lohmann DR, Brandt B, Höpping W, Passarge E, Horsthemke B. The spectrum of RB1 germ-line mutations in hereditary retinoblastoma. *Am J Hum Genet.* 1996;58:940–9.
20. Antonarakis ES, Lu C, Wang H, Lubner B, Nakazawa M, Roeser JC, et al. AR-V7 and resistance to enzalutamide and abiraterone in prostate cancer. *N Engl J Med.* 2014;371:1028–38.
21. Cho WJ, Oliveira DSM, Najy AJ, Mainetti LE, Aoun HD, Cher ML, et al. Gene expression analysis of bone metastasis and circulating tumor cells from metastatic castrate-resistant prostate cancer patients. *J Transl Med.* 2016;14:72.
22. Robinson D, Van Allen EM, Wu Y-M, Schultz N, Lonigro RJ, Mosquera J-M, et al. Integrative clinical genomics of advanced prostate cancer. *Cell.* 2015;161:1215–28.
23. Beltran H, Prandi D, Mosquera JM, Benelli M, Puca L, Cyrta J, et al. Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. *Nat Med.* 2016;22:298–305.
24. Zhou Z, Flesken-Nikitin A, Corney DC, Wang W, Goodrich DW, Roy-Burman P, et al. Synergy of p53 and Rb deficiency in a conditional mouse model for metastatic prostate cancer. *Cancer Res.* 2006;66:7889–98.
25. Tan H-L, Sood A, Rahimi HA, Wang W, Gupta N, Hicks J, et al. Rb loss is characteristic of prostatic small cell neuroendocrine carcinoma. *Clin Cancer Res.* 2014;20:890–903.
26. Massard C, Oulhen M, Le Moulec S, Auger N, Foulon S, Abou-Lovergne A, et al. Phenotypic and genetic heterogeneity of tumor tissue and circulating tumor cells in patients with metastatic castration resistant prostate cancer: a report from the PETRUS prospective study. *Oncotarget.* 2016;7:55069–82.
27. Hosono S, Faruqi AF, Dean FB, Du Y, Sun Z, Wu X, et al. Unbiased whole-genome amplification directly from clinical samples. *Genome Res.* 2003;13:954–64.
28. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol.* 2011;29:24–6.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit



Journal of the National Cancer Institute

Genomic heterogeneity within individual prostate cancer foci impacts predictive biomarkers of targeted therapy --Manuscript Draft--

Manuscript Number:	JNCI-17-1318
Full Title:	Genomic heterogeneity within individual prostate cancer foci impacts predictive biomarkers of targeted therapy
Article Type:	Article
Corresponding Author:	David VanderWeele, MD PhD National Cancer Institute Bethesda, Maryland UNITED STATES
Order of Authors Secondary Information:	
Keywords:	Clonal evolution; Copy number alteration; Exome sequencing; Genomic heterogeneity; Multiregion sequencing; Predictive biomarkers; Prostate cancer; Subclonal architecture; targeted therapy
Section/Category:	Tumor Biology/Physiology
Suggested Reviewers:	
Opposed Reviewers:	
Manuscript Classifications:	Genetics; Prostate Cancer; Targeted Therapies
Author Comments:	<p>September 7, 2017</p> <p>Dear Dr. Ganz,</p> <p>We would like to submit the manuscript "Genomic heterogeneity within individual prostate cancer foci impacts predictive biomarkers of targeted therapy" for consideration for publication as an Article in JNCI.</p> <p>Most lethal prostate cancer progresses from relapse of primary disease. Given known heterogeneity between multiple foci, prognosis is generally attributed to a dominant index focus. The heterogeneity within the index focus of aggressive, non-metastatic disease, however, is not known.</p> <p>We use multiregion genomic analysis to examine the heterogeneity of potentially lethal, treatment-naïve prostate cancer. Our data demonstrate several significant and clinically important findings. First, there is marked genomic heterogeneity in high risk, treatment-naïve disease, indicating the heterogeneity found in metastatic castrate resistant prostate cancer (mCRPC) can also be found at much earlier stages. Second, this heterogeneity extends to potential genetic biomarkers, such as defects in homologous recombination repair or alterations in the PI3K pathway. Biomarkers present in one region are often discordant with other regions, and biomarkers identified in the prostate tumor are discordant with synchronous metastases. Third, there are numerous co-occurring alterations in pathways typically associated with mCRPC, though found in subclonal populations in treatment-naïve disease. Fourth, there is widespread dispersal of subclonal populations, so even widespread alterations are</p>

	<p>often subclonal.</p> <p>Currently there is uncertainty about how to manage high risk, non-metastatic disease, in part due to uncertainty about its biology. The LATITUDE and STAMPEDE trials are practice changing to include abiraterone with prednisone for castrate sensitive disease with extensive metastases. Physicians are less comfortable extending this management to high-risk, non-metastatic disease, though these patients were included in the STAMPEDE trial. Though there is clearly a failure-free survival benefit for non-metastatic disease, the overall survival data is not mature. Moreover, there is uncertainty regarding the biology of nonmetastatic disease relative to metastatic disease.</p> <p>The data in this study are the most thorough genomic characterization to date of potentially lethal, nonmetastatic, treatment-naïve disease. While the findings presented here cause concern about the potential for resistance to systemic therapy that is associated with such marked genetic heterogeneity, they also demonstrate that the biology of aggressive, non-metastatic disease is similar to metastatic, therapy-resistant disease. Given that biological similarity, these data give the rationale for the patient population of the STAMPEDE trial, pooling high risk localized disease (as included in the present study) with widely metastatic disease. And they lend support to embrace treatment of non-metastatic disease in a similar fashion to widely metastatic disease.</p> <p>Though PARP inhibitors and Akt inhibitors are not yet approved for prostate cancer, they are the targeted therapies on the horizon for mCRPC. One response to the CHAARTED, STAMPEDE, and LATITUDE trials is the desire to move additional systemic therapies to an earlier disease state. The data presented here also provide insight into using biomarkers from intraprostatic disease to guide use of those therapies.</p> <p>Our manuscript has significant implications for the tens of thousands of patients diagnosed with potentially lethal, aggressive primary prostate cancer every year. We feel this timely manuscript will be welcomed by the prostate cancer community as it deliberates about how best to manage high risk, treatment-naïve disease. We hope you agree.</p> <p>Thank you for your consideration of this work for publication in JNCI. Please feel free to contact me directly should you have any questions.</p> <p>Kind regards,</p> <p>David J. VanderWeele MD, PhD Assistant Clinical Investigator Laboratory for Genitourinary Cancer Pathogenesis LGCP, NCI, NIH 37 Convent Drive, Rm 1066A Bethesda, MD 20892 301.443.2622 david.vanderweele@nih.gov</p>
Additional Information:	
Question	Response
Choose one statement regarding the planning and drafting of this manuscript:	All authors of this research paper have directly participated in the planning, execution, or analysis of the study.
Enter your initials in the box to confirm the following statement: All authors of this paper have read and approved the final	DVW

version submitted.	
Enter your initials in the box to confirm the following statement: The contents of this manuscript have not been copyrighted or published previously.	DVW
Enter your initials in the box to confirm the following statement: The contents of this manuscript are not now under consideration for publication elsewhere.	DVW
Enter your initials in the box to confirm the following statement: The contents of this manuscript will not be copyrighted, submitted, or published elsewhere while acceptance by the Journal is under consideration.	DVW
Choose one statement regarding manuscripts directly related to this submission:	There are no directly related manuscripts or abstracts, published or unpublished, by any author(s) of this paper.
Does the research described in this manuscript meet ethical guidelines, including adherence to the legal requirements of the study country?	Yes
Is this submission one of a set of linked or companion papers, which should be considered for publication together?	No
Excluding Supplementary Tables, how many tables are included in your manuscript?	1
Excluding Supplementary Figures, how many figures are included in your manuscript?	4
Does this manuscript submission contain supplemental materials to be published online?	Yes
<p>Please note the following limits on the total number of supplemental tables plus supplemental figures:</p> <p>Articles, Reviews, Systematic Reviews, and Meta-Analyses: 8</p> <p>Commentaries and Mini-Reviews: 4</p> <p>Brief Communications: 2</p> <p>Correspondence, Responses, and Editorials: 1</p> <p>Does this submission exceed these limits? If so, select Yes and provide a justification in the resulting box.</p> <p>as follow-up to "Does this manuscript submission contain supplemental materials to be published online?"</p>	No
<p>Please check the box below to confirm the following statement:</p> <p><input type="checkbox"/> I hereby grant to Oxford University Press a non-exclusive license for the duration of the copyright period to publish the data</p>	I confirm this statement.

<p>supplement/supplementary materials in all languages and media. I also grant to Oxford University Press the right to grant third party permissions to republish the data supplement/supplementary materials in whole or parts thereof in any medium without limitation. As the author, I retain the copyright and all other rights in the data supplement/supplementary materials.</p> <p>as follow-up to "Does this manuscript submission contain supplemental materials to be published online?"</p>	
Please indicate which of the following standard reporting guidelines were used for this submission:	Not Applicable
If you selected "Other" in the previous question, which reporting guideline did you use for this submission?	
Is this a solicited manuscript?	No
I confirm that authors of this paper agree to be accountable for all aspects of the work, such that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.	I confirm this statement.
Order of Authors:	David J. VanderWeele, MD PhD
	Richard Finney
	Kotoe Katayama
	Marc Gillard
	Gladell Paner
	Seiya Imoto
	Rui Yamaguchi
	David Wheeler
	Justin Lack
	Maggie Cam
	Andrea Pontier
	Yen Min Thi Nguyen
	Kazuhiro Maejima
	Aya Sasaki-Oku
	Kaoru Nakano
	Hiroko Tanaka
	Donald Vander Griend
	Michiaki Kubo
	Mark J. Ratain
	Satoru Miyano
	Hidewaki Nakagawa



Center for
Cancer Research

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
National Institutes of Health

September 7, 2017

Dear Dr. Ganz,

We would like to submit the manuscript “Genomic heterogeneity within individual prostate cancer foci impacts predictive biomarkers of targeted therapy” for consideration for publication as an Article in *JNCI*.

Most lethal prostate cancer progresses from relapse of primary disease. Given known heterogeneity between multiple foci, prognosis is generally attributed to a dominant index focus. The heterogeneity within the index focus of aggressive, non-metastatic disease, however, is not known.

We use multiregion genomic analysis to examine the heterogeneity of potentially lethal, treatment-naïve prostate cancer. Our data demonstrate several significant and clinically important findings. First, there is marked genomic heterogeneity in high risk, treatment-naïve disease, indicating the heterogeneity found in metastatic castrate resistant prostate cancer (mCRPC) can also be found at much earlier stages. Second, this heterogeneity extends to potential genetic biomarkers, such as defects in homologous recombination repair or alterations in the PI3K pathway. Biomarkers present in one region are often discordant with other regions, and biomarkers identified in the prostate tumor are discordant with synchronous metastases. Third, there are numerous co-occurring alterations in pathways typically associated with mCRPC, though found in subclonal populations in treatment-naïve disease. Fourth, there is widespread dispersal of subclonal populations, so even widespread alterations are often subclonal.

Currently there is uncertainty about how to manage high risk, non-metastatic disease, in part due to uncertainty about its biology. The LATITUDE and STAMPEDE trials are practice changing to include abiraterone with prednisone for castrate sensitive disease with extensive metastases. Physicians are less comfortable extending this management to high-risk, non-metastatic disease, though these patients were included in the STAMPEDE trial. Though there is clearly a failure-free survival benefit for non-metastatic disease, the overall survival data is not mature. Moreover, there is uncertainty regarding the biology of nonmetastatic disease relative to metastatic disease.

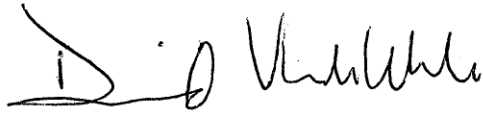
The data in this study are the most thorough genomic characterization to date of potentially lethal, nonmetastatic, treatment-naïve disease. While the findings presented here cause concern about the potential for resistance to systemic therapy that is associated with such marked genetic heterogeneity, they also demonstrate that the biology of aggressive, non-metastatic disease is similar to metastatic, therapy-resistant disease. Given that biological similarity, these data give the rationale for the patient population of the STAMPEDE trial, pooling high risk localized disease (as included in the present study) with widely metastatic disease. And they lend support to embrace treatment of non-metastatic disease in a similar fashion to widely metastatic disease.

Though PARP inhibitors and Akt inhibitors are not yet approved for prostate cancer, they are the targeted therapies on the horizon for mCRPC. One response to the CHAARTED, STAMPEDE, and LATITUDE trials is the desire to move additional systemic therapies to an earlier disease state. The data presented here also provide insight into using biomarkers from intraprostatic disease to guide use of those therapies.

Our manuscript has significant implications for the tens of thousands of patients diagnosed with potentially lethal, aggressive primary prostate cancer every year. We feel this timely manuscript will be welcomed by the prostate cancer community as it deliberates about how best to manage high risk, treatment-naïve disease. We hope you agree.

Thank you for your consideration of this work for publication in *JNCI*. Please feel free to contact me directly should you have any questions.

Kind regards,

A handwritten signature in black ink, appearing to read 'David J. VanderWeele', with a stylized, cursive script.

David J. VanderWeele MD, PhD
Assistant Clinical Investigator
Laboratory for Genitourinary Cancer Pathogenesis
LGCP, NCI, NIH
37 Convent Drive, Rm 1066A
Bethesda, MD 20892
240.760.6829
david.vanderweele@nih.gov

Title:**Genomic heterogeneity within individual prostate cancer foci impacts
predictive biomarkers of targeted therapy**

Authors: David J. VanderWeele MD PhD^{1,2}, Richard Finney³, Kotoe Katayama PhD⁴, Marc Gillard PhD^{2,5}, Gladell Paner MD⁶, Seiya Imoto⁴, Rui Yamaguchi⁴, David Wheeler PhD³, Justin Lack PhD³, Maggie Cam PhD³, Andrea Pontier MS², Yen Min Thi Nguyen¹, Kazuhiro Maejima⁷, Aya Sasaki-Oku⁷, Kaoru Nakano⁷, Hiroko Tanaka⁴, Donald Vander Griend PhD⁵, Michiaki Kubo MD PhD⁷, Mark J. Ratain MD², Satoru Miyano PhD⁴, Hidewaki Nakagawa MD PhD⁷

Affiliations:

¹Laboratory for Genitourinary Pathogenesis, Center for Cancer Research, National Cancer Institute, Bethesda, MD, 20892 USA

²Department of Medicine, University of Chicago, Chicago, IL, 60615 USA

³Center for Cancer Research Collaborative Bioinformatics Resource, Center for Cancer Research, National Cancer Institute, Bethesda, MD, 20892 USA

⁴Human Genome Center, Institute of Medical Science, University of Tokyo, Minato-ku, Tokyo, 108-8639 Japan

⁵Department of Surgery, University of Chicago, Chicago, IL, 60615 USA

⁶Department of Pathology, University of Chicago, Chicago, IL, 60615 USA

⁷Laboratory for Genome Sequencing Analysis, RIKEN Center for Integrative Medical Sciences,
RIKEN, Yokohama City, Kanagawa, 230-0045 Japan

Corresponding author

David VanderWeele MD PhD

Assistant Clinical Investigator

Laboratory for Genitourinary Cancer Pathogenesis

National Cancer Institute

37 Convent Drive, Rm 1066A

Bethesda, MD 20892

240.760.6829 (Ph)

240.541.4503 (F)

david.vanderweele@nih.gov

ABSTRACT

Background: Most lethal prostate cancer progresses from relapse of aggressive primary disease. Given known heterogeneity between multiple foci, prognosis is generally attributed to a dominant index focus. Targeted therapies show promise for aggressive prostate cancer, and predicted response is based on the presence of genomic biomarkers. The present study characterizes the genomic heterogeneity, including that of predictive biomarkers, within the index focus of aggressive primary prostate cancer.

Methods: Genomic characterization was performed on 70 regions from the index foci of ten patients with treatment-naïve, aggressive primary prostate cancer. Five to nine regions of each index focus were evaluated by exome and low coverage whole genome sequencing. Intrafocal genomic heterogeneity and heterogeneity of alterations that predict response to therapy was determined.

Results: Exome sequencing and copy number estimates demonstrate branched evolution with >75% of point mutations being subclonal, including numerous pathways associated with castrate resistant prostate cancer. There is wide dispersal of comingled subclones, with individual subclonal populations detectable across all regions sampled. Intraindividual heterogeneous alterations in 5 genes predict response to targeted therapies. Within biomarker-positive cases, 25% of intraprostatic regions are biomarker-negative, with discordance between intraprostatic regions and lymph node metastases.

Conclusions: Treatment-naïve, nonmetastatic prostate cancer has marked intrafocal heterogeneity. Numerous alterations in pathways associated with castrate resistant prostate cancer are present in subclonal populations, including biomarkers predictive of response to targeted therapy.

INTRODUCTION

Thirty percent of patients receiving definitive therapy for newly diagnosed prostate cancer relapse, with higher rates of relapse for those with high risk or locally advanced disease ¹⁻³. For over seven decades, systemic therapy has focused on the androgen receptor (AR), a disease-specific target. One goal of genomic characterizations of both untreated and advanced prostate cancer is to identify subtypes of disease that predict sensitivity to new therapies ⁴⁻⁶. These studies have suggested that in advanced disease over 60% of patients have targetable altered pathways in addition to the AR signaling pathway ⁶. The PARP inhibitor olaparib is the targeted therapy furthest in clinical development, for patients with tumors harboring mutations in DNA repair pathways, comprising approximately 20% of advanced disease ⁶⁻⁸. Recently ipatasertib has also demonstrated efficacy in patients with loss of the tumor suppressor gene PTEN ⁹. It is hoped that further efforts will define additional subsets of disease with specific susceptibilities, effectively dividing the diagnosis into several different molecularly-based diagnoses.

The efficacy and durability of response to targeted therapies depends in part on the prevalence of the targeted alteration within the tumor cell population. Multifocal prostate cancer is known to have early divergence between foci, but there is often a dominant, high grade index focus considered to represent clinically significant disease ¹⁰⁻¹⁴. Supporting this, there is relatively little inpatient heterogeneity in metastatic disease when prostate cancer becomes lethal ^{15,16}. However, the extent of heterogeneity within the index focus of aggressive prostate cancer is unknown.

We undertook multiregion genomic evaluation of an individual index focus from ten patients with localized (lymph node negative) or locally advanced (lymph node positive) disease, representing those most likely to relapse. To our knowledge, this is the first genomic evaluation focused on potentially lethal, locally advanced prostate cancer. The present study describes branched genomic evolution characterized by marked heterogeneity in both localized and locally advanced disease. Compared to localized disease, locally advanced disease is characterized by increased copy number changes and numerous subclonal alterations in pathways associated with castrate resistant prostate cancer (CRPC). Significantly, alterations predictive of response to targeted therapy are heterogeneous within the prostate and between intraprostatic disease and lymph node metastatic disease.

METHODS

Subjects and consent

Prostatectomy specimens were collected from subjects enrolled in “Prostate Cancer Sample Collection” tissue collection protocol of the University of Chicago Medical Center (see **Supplemental Methods**). All individuals included in the study provided written informed consent. All cases were reviewed by a genitourinary pathologist (GP), and 3D reconstructions of the index focus were created based on location and histologic appearance of tumor within each block.

Genomic analysis

Germline DNA was extracted from blood using a DNeasy kit (Qiagen), except for case 1027, when it was obtained from histologically benign prostate tissue. DNA was extracted from

tumor regions using PAXgene Tissue DNA kit (Qiagen) and exome and low coverage whole genome sequencing performed (see **Supplemental Methods**). A subset of variants was verified with ultradeep targeted resequencing on a MiSeq (**Supplemental Table 1**). Subsequently seventeen regions were excluded from further analysis due to low tumor cell purity (<30%) based on variant allele frequency of identified mutations leaving 70 regions with sufficient sequencing quality and tumor cell purity. Positions harboring mutations that passed quality filters were then evaluated for all regions in the corresponding case using samtools mpileup. Mutations were considered present if the variant allele frequency (VAF) > 1%, and VAF > 10% in at least one region. Indels were confirmed using IGV.

Copy number measures were generated by BICseq2¹⁷ (see **Supplemental Methods**) setting filters to p-value < 0.01 and log2 ratio < -0.2 or > 0.2¹⁷. Whole genome sequencing data was unavailable for case 1009, and copy number alterations were determined using CNVkit¹⁸ where log2 ratio was < -0.2 or > 0.2 and confidence interval excluded one. To avoid falsely elevating the fraction of subclonal events due to lower tumor purity in some samples, if log2 ratios in all regions from the case were in the same direction (>0 or <0), an alteration was considered clonal. If not, it was considered subclonal.

Immunohistochemistry was performed for ERG expression. Nuclear ERG expression intensity and staining distribution were scored on a scale of 0-3 by a genitourinary pathologist (GP).

3D representations of prostatectomy specimens and network trees

Representations of prostatectomy specimens were created using Povray based on manual mapping of the tumor focus and regions sampled. Network analysis was constructed using Network ^{19,20} from features comprising all somatic mutations identified.

Mutations per cell

The average number of mutation per cell was calculated from exome sequencing data based on Landau et al ²¹ (see **Supplemental Methods**). Average number of mutations per cell was calculated as $(\text{Mutation VAF}) * (\text{purity} * \text{ploidy} + 2 * (1 - \text{purity})) / \text{purity}$.

Predictive genetic biomarkers

Potential biomarkers were queried using The Drug Gene Interaction Database (dgidb.genome.wustl.edu) ²². A biomarker was considered positive if it contained both a nonsynonymous mutation and deletion ($\log_2 < -0.2$) or a deep deletion ($\log_2 < -0.4$). To avoid falsely elevating heterogeneity of biomarkers a deletion of \log_2 ratio < -0.2 was also considered a biomarker without mutation if at least one region in the case had deep deletion at that locus.

Statistics

A two-sided student's t-test using R was used to analyze the difference between the number of mutations and the percent genome altered in localized versus locally advanced cases. A log-rank Mantel-Cox test was performed using Prism to evaluate the difference in biochemical recurrence between cases with more or less than the median CRPC-associated alterations.

RESULTS

Intrafocal heterogeneity of aggressive primary prostate cancer

To examine the heterogeneity of treatment-naïve, aggressive primary prostate cancer, multiregion sequencing was performed on an individual index focus from the prostatectomy specimen of ten patients (**Supplemental Figure 1**). Clinical characteristics are typical of patients with aggressive disease, with Gleason scores 7-9 and pathologic tumor stage pT2-T3 (**Table 1**). Five cases are localized (lymph node negative), five are locally advanced (lymph node positive). Three patients had biochemical recurrence following prostatectomy, and four underwent additional therapy after prostatectomy. For each case, exome and low coverage whole genome sequencing was performed on five to nine regions of a single contiguous index focus. Lymph node metastases obtained at the time of prostatectomy and seminal vesicle regions involved through direct extension were included for two cases each (1010 and 1022; 1015 and 1022, respectively).

We identified 887 unique mutation events across 70 index focus regions (88.7 variants/case, ~1.3 nonsilent variants/Mb) (**Supplemental Table 2**), with a higher number of mutations in locally advanced than in localized cases (108.6 vs 68.6, respectively). A subset of mutations was verified with ultra-high coverage (~40,000x) targeted resequencing. Of 107 mutations assessed, 2 (2%) were not confirmed, and none was found in germline (**Supplemental Table 1**). Allele frequencies measured by exome sequencing were similar to those measured by targeted resequencing ($R^2=0.861$, **Supplemental Figure 2**).

Exome sequencing demonstrated marked heterogeneity of single nucleotide mutations and indels within the cancer genome of the index focus in all cases examined (**Figure 1A**). In

each focus trunk mutations (mutations identified in all regions) comprised the minority.

Alterations in ERG, SPOP, FOXA1, IDH1 are among the drivers proposed to define prostate cancer subtypes in 75% of cases ⁵. Five of 10 cases presented here could be assigned one of these subtypes (3 ERG overexpression, 1 FOXA1 mutation, 1 SPOP mutation) (**Supplemental Figure 3**). As with a single region sequencing approach, subtypes were mutually exclusive. As expected, these alterations occurred early, with the alteration identified in every region sampled in all but one case (ERG in case 1029).

Stability of mutation processes over time

Several studies have presented lines of evidence for a field effect in prostate cancer ^{13,23}, suggesting there is a long accumulation of mutations over the life of the patient. Given this, we examined the types of alterations contributing to early (trunk), late (leaf), and middle (branch) events. In contrast to the 60% or more clonal mutations shown for 10 other solid tumors ²⁴, fewer than 25% of variants identified were trunk variants. Despite the numerically increased number of mutations in locally advanced cancer, there was no increase in fraction of late mutations in these cases (**Supplemental Figure 4A**). Unlike other malignancies ^{24–26}, we also found no evidence for changes in mutational processes over time, with similar patterns of transitions vs transversions and patterns of non-synonymous, synonymous, premature stop, and splice site mutations among trunk, branch, and leaf mutations (**Supplemental Figure 4B,C**).

Heterogeneity among structural changes

Chromosomal events such as TMPRSS2:ERG fusion appear to be early events in prostate cancer, present even in a fraction of premalignant lesions ^{27,28}. To identify the regional pattern of

chromosomal changes, we performed copy number estimates from low-coverage whole genome sequencing (**Figure 1B, Supplemental Table 3**). There were many canonical copy number changes that were estimated, including high frequency of loss of chromosomes 8p, 13, and 16, focal deletion on chromosome 10, and gain of chromosomes 8q and 7. Locally advanced cases had higher mean percent genome altered than those without lymph node involvement, which is associated with a worse prognosis^{29,30} (**Figure 1C**). Interestingly, two of the four cases harboring regions with the highest percent of the genome altered (> 0.2) also contained regions with low copy number burden (~ 0.05), suggesting there is focal genome instability.

Correlation between histology and genetic features

We constructed network trees for each case based on mutation patterns. These were overlaid on 3D representations of the prostatectomy specimens (**Figure 2, Supplemental Figure 5**). In cases with histologic heterogeneity, it appeared the relationship between regions was as likely to correlate with histologic similarity than spatial proximity. In case 1010 for example, despite their proximity to regions C36A and C46A, respectively, regions C36B and C46B share high grade features and late divergence of their cancer genome, supported by shared mutations and copy number changes. Both are more similar histologically and diverge later from C50A than from C36A and C46A.

We also observed early divergence of spatially proximal regions without obvious correlations with histology (early divergence of region B3A in case 1024, **Supplemental Figure 5**). In most cases, however, there was consistent histologic appearance across the focus

(**Supplemental Figure 1**), and spatially proximal regions showed late divergence. This was most clearly seen in cases 1024 and 1034 (**Supplemental Figure 5**).

Subclonal alterations in clinically significant pathways

A recent report of the genomic landscape of castrate resistant prostate cancer (CRPC) demonstrated alterations in statistically or clinically significant genes from several pathways that contribute to prostate cancer progression, including androgen receptor, cell cycle, PI3K, RAF/RAS, WNT, DNA repair, and chromatin modifier pathways ⁶. Many of these genes are found to be altered at low frequency in treatment-naïve disease ⁵. We hypothesized these genes might be altered in more cases of treatment-naïve disease than previously appreciated, but in subclonal populations below the level of detection in other studies. Indeed, all ten cases have alterations in at least two of 41 CRPC-related genes (**Figure 3A**). Moreover, those with higher than median subclonal alterations in CRPC pathways were more likely to have recurrence than those below the median (p-value < 0.05, log-rank test) (**Figure 3B**). The majority of the alterations we identified were chromosomal events, highlighting the frequency and importance of chromosome level events compared to single nucleotide mutations ³¹.

Given that many of the CRPC pathways altered are potential therapeutic targets, we evaluated the heterogeneity in genetic biomarkers that predict response to targeted therapy. Using concurrent copy loss and mutation or deep deletion we identified five genes that predict susceptibility to two therapies with demonstrated efficacy in CRPC ^{8,9}, with individual cases harboring biomarker alterations in up to three separate genes (**Figure 3C**). One case (1009) had germline mutation in the DNA-binding domain of BRCA2 that is of unknown significance

(p.Ala2730Pro), with concurrent somatic copy loss of BRCA2. The remaining predictive biomarkers were heterogeneous among regions sampled such that, of biomarker-positive cases, 25% of regions were biomarker-negative. Moreover, there was discordance between intraprostatic regions and lymph node metastases: in two cases with lymph node metastases available, only one of three biomarker alterations found in intraprostatic regions was also found in the lymph node, and one of two alterations identified in a lymph node was found in the prostate (**Figure 3C**).

Spatial dispersion of subclones

For a targeted therapy to be effective, the targeted alteration ought to be found in every cancer cell. We assume that mutations detected in every region of a focus (trunk mutations) occur early. It is not clear, however, if these mutations are detected in every region because they are found in every cell, or because they are found in subclones that have dispersed widely. We examined the variant allele frequencies (VAFs) of mutations in coincident regions to determine if wide spatial dispersion indicates it is indeed an early event. As expected, trunk mutations are found at high allele frequency, generally correlating with tumor cell purity (typically ~0.30, representing ~60% tumor cell purity; **Figure 4A**). Leaf mutations, detected in a single region, have low allele frequency within that region. Most branch mutations also have low allele frequency within a region, indicating subclones can have wide (present in multiple regions) but sparse (low allele frequency) dispersion.

To evaluate this further, we generated 2D density plots for representative spatially distinct regions, representing VAF normalized for tumor cell purity and ploidy (akin to cancer

cell fraction) (**Figure 4B,C**). This demonstrates that branch subclone populations have highly variable cancer cell fractions, consistent with a selection mode of tumor evolution ³². Notably, even mutations identified in all regions (trunk mutations) can have low allele frequency within a region, indicating they are found in a widely dispersed subclone rather than being truly ubiquitous.

DISCUSSION

We present here the largest multiregion genomic evaluation of the index focus of aggressive primary prostate cancer. This study demonstrates the heterogeneity of treatment-naive, aggressive disease, with fewer than 25% of mutations found in all regions of the focus. Coincident foci are known to be divergent, leading to reliance on the dominant index focus for prognostication. These data demonstrate remarkable heterogeneity within the clinically relevant index focus.

The treatment of prostate cancer has recently entered the genomic era, with PARP inhibitor therapy based on alterations in DNA repair pathways ⁶⁻⁸ demonstrating response in 88% of biomarker positive patients, and an Akt inhibitor has also demonstrated benefit in biomarker-positive patients ⁹. There is interest in moving these therapies up to an earlier disease state, with trials underway testing olaparib in the neoadjuvant or biochemically recurrent space (NCT02324998, NCT03047135), where biomarker evaluation is presumably based on the index lesion of the primary tumor. In the present study we show that while prostate driver alterations are early events, predictive biomarkers such as ATM, BRCA1, and BRCA2 can be altered in a small fraction of cells of the index focus. Moreover, we find that even wide spatial dispersal of

an alteration does not imply it is ubiquitous. In these cases, one would expect a transient response to targeted therapy. This is in line with studies demonstrating that intratumoral heterogeneity is linked with poor outcome and therapy resistance^{25,33–35}.

Currently there is uncertainty about how to manage high risk, non-metastatic disease, in part due to uncertainty about its biology. The LATITUDE and CHAARTED trials^{36,37} indicate a remarkable survival benefit in moving life-prolonging therapies up to an earlier disease state, and the STAMPEDE trial^{38,39} suggests this also applies to non-metastatic disease, though the survival data is less mature. The findings presented here indicate that aggressive, non-metastatic disease, especially locally advanced disease, is remarkably similar to metastatic, castrate-resistant disease. It harbors a high percent of the genome altered by copy number alterations, and there are numerous alterations typically associated with metastatic CRPC.

Our data expands the growing body of work demonstrating heterogeneity of primary prostate cancer. Previous work demonstrated interfocal^{10,11,40} or intra- and interfocal heterogeneity^{12,13,41}, though these have focused on lower risk, lower stage disease, a disease state that is likely to be cured by surgery and is more dissimilar to widely metastatic disease. Multiple expression-based assays have been developed to refine prognosis of localized disease, and Wei et al⁴¹ demonstrated intra- and to a greater extent interfocal heterogeneity in these assays. Two studies^{42,43} have compared primary and metastatic disease and demonstrated cases of biomarker-negative prostate tumors and biomarker-positive metastases due to the accumulation of additional alterations in metastatic disease, which would presumably lead to a

false negative primary tissue biomarker. The present study now demonstrates a positive prostate tissue biomarker not representative of all tumor cells, predicted to lead to poor response.

These data demonstrate there can be marked heterogeneity among regions of an individual index focus in aggressive primary prostate cancer. This heterogeneity also applies to genetic biomarkers, highlighting the risk of selecting biomarker-driven therapies based on primary tissue. The concern about the heterogeneity of CRPC has led to the proposal of novel therapeutic strategies to compensate for this ⁴⁴. Our data suggest concerns about tumor heterogeneity also apply to the management of androgen sensitive disease.

Financial support:

This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs, through the Prostate Cancer Research Program (grant number W81XWH-13-1-0451 (DVW)). Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the Department of Defense. The U.S. Army Medical Research Acquisition Activity, 820 Chandler Street, Fort Detrick MD 21702-5014 is the awarding and administering acquisition office. This work was also supported by the University of Chicago Cancer Center Support Grant (grant number P30 CA014599), and the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. Exome sequencing analysis was performed in the super-computing resource “SHIROKANE” in Human Genome Center, The University of Tokyo. This work also utilized the computational resources of the NIH HPC Biowulf cluster (<http://hpc.nih.gov>).

Acknowledgements: The authors thank the patients who participated in this study, as well as the University of Chicago Genomics Facility and Human Tissue Research Center, and the NCI Center for Cancer Research Collaborative Bioinformatics Resource for their assistance.

References

1. Mullins JK, Feng Z, Trock BJ, Epstein JI, Walsh PC, Loeb S. The Impact of Anatomical Radical Retropubic Prostatectomy on Cancer Control: The 30-Year Anniversary. *J Urol*. 2012;188(6):2219-2224. doi:10.1016/j.juro.2012.08.028.
2. Boorjian SA, Eastham JA, Graefen M, et al. A Critical Analysis of the Long-Term Impact of Radical Prostatectomy on Cancer Control and Function Outcomes. *Eur Urol*. 2012;61(4):664-675. doi:10.1016/j.eururo.2011.11.053.
3. Morgan TM, Meng MV, Cooperberg MR, et al. A risk-adjusted definition of biochemical recurrence after radical prostatectomy. *Prostate Cancer Prostatic Dis*. 2014;17(2):174-179. doi:10.1038/pcan.2014.5.
4. Taylor BS, Schultz N, Hieronymus H, et al. Integrative genomic profiling of human prostate cancer. *Cancer Cell*. 2010;18(1):11-22. doi:10.1016/j.ccr.2010.05.026.
5. Abeshouse A, Ahn J, Akbani R, et al. The Molecular Taxonomy of Primary Prostate Cancer. *Cell*. 2015;163(4):1011-1025. doi:10.1016/j.cell.2015.10.025.
6. Robinson D, Van Allen EM, Wu Y-M, et al. Integrative clinical genomics of advanced prostate cancer. *Cell*. 2015;161(5):1215-1228. doi:10.1016/j.cell.2015.05.001.
7. VanderWeele DJ, Paner GP, Fleming GF, Szmulewitz RZ. Sustained Complete Response to Cytotoxic Therapy and the PARP Inhibitor Veliparib in Metastatic Castration-Resistant Prostate Cancer - A Case Report. *Front Oncol*. 2015;5:169. doi:10.3389/fonc.2015.00169.
8. Mateo J, Carreira S, Sandhu S, et al. DNA-Repair Defects and Olaparib in Metastatic Prostate Cancer. *N Engl J Med*. 2015;373(18):1697-1708. doi:10.1056/NEJMoa1506859.
9. de Bono JS, De Giorgi U, Massard C, et al. PTEN loss as a predictive biomarker for the Akt inhibitor ipatasertib combined with abiraterone acetate in patients with metastatic castration-resistant prostate cancer (mCRPC). *Ann Oncol*. 2016;27(suppl_6). doi:10.1093/annonc/mdw372.02.
10. VanderWeele DJ, Brown CD, Taxy JB, et al. Low-grade prostate cancer diverges early from high grade and metastatic disease. *Cancer Sci*. 2014;105(8):1079-1085. doi:10.1111/cas.12460.
11. Lindberg J, Klevebring D, Liu W, et al. Exome sequencing of prostate cancer supports the hypothesis of independent tumour origins. *Eur Urol*. 2013;63(2):347-353. doi:10.1016/j.eururo.2012.03.050.
12. Boutros PC, Fraser M, Harding NJ, et al. Spatial genomic heterogeneity within localized, multifocal prostate cancer. *Nat Genet*. 2015;47(7):736-745. doi:10.1038/ng.3315.
13. Cooper CS, Eeles R, Wedge DC, et al. Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and

- morphologically normal prostate tissue. *Nat Genet.* 2015;47(4):367-372. doi:10.1038/ng.3221.
14. Wei L, Wang J, Lampert E, et al. Intratumoral and Intertumoral Genomic Heterogeneity of Multifocal Localized Prostate Cancer Impacts Molecular Classifications and Genomic Prognosticators. *Eur Urol.* doi:10.1016/j.eururo.2016.07.008.
 15. Aryee MJ, Liu W, Engelmann JC, et al. DNA methylation alterations exhibit intraindividual stability and interindividual heterogeneity in prostate cancer metastases. *Sci Transl Med.* 2013;5(169):169ra10. doi:10.1126/scitranslmed.3005211.
 16. Kumar A, Coleman I, Morrissey C, et al. Substantial interindividual and limited intraindividual genomic diversity among tumors from men with metastatic prostate cancer. *Nat Med.* 2016;22(4):369-378. doi:10.1038/nm.4053.
 17. Xi R, Lee S, Xia Y, Kim T-M, Park PJ. Copy number analysis of whole-genome data using BIC-seq2 and its application to detection of cancer susceptibility variants. *Nucleic Acids Res.* 2016;44(13):6274-6286. doi:10.1093/nar/gkw491.
 18. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLOS Comput Biol.* 2016;12(4):e1004873. doi:10.1371/journal.pcbi.1004873.
 19. Network FAQ. <http://www.fluxus-engineering.com/sharepub.htm#a1>. Accessed May 26, 2016.
 20. Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol.* 1999;16(1):37-48.
 21. Landau DA, Carter SL, Getz G, Wu CJ. Clonal evolution in hematological malignancies and therapeutic implications. *Leukemia.* 2014;28(1):34-43. doi:10.1038/leu.2013.248.
 22. Griffith M, Griffith OL, Coffman AC, et al. DGIdb: mining the druggable genome. *Nat Methods.* 2013;10(12):1209-1210. doi:10.1038/nmeth.2689.
 23. Gaisa NT, Graham TA, McDonald SA, et al. Clonal architecture of human prostatic epithelium in benign and malignant conditions. *J Pathol.* 2011;225(2):172-180. doi:10.1002/path.2959.
 24. de Bruin EC, McGranahan N, Mitter R, et al. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science.* 2014;346(6206):251-256. doi:10.1126/science.1253462.
 25. Zhang J, Fujimoto J, Zhang J, et al. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science.* 2014;346(6206):256-259. doi:10.1126/science.1256930.

26. Gerlinger M, Horswell S, Larkin J, et al. Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nat Genet.* 2014;46(3):225-233. doi:10.1038/ng.2891.
27. Cerveira N, Ribeiro FR, Peixoto A, et al. TMPRSS2-ERG gene fusion causing ERG overexpression precedes chromosome copy number changes in prostate carcinomas and paired HGPIN lesions. *Neoplasia N Y N.* 2006;8(10):826-832. doi:10.1593/neo.06427.
28. Morais CL, Guedes LB, Hicks J, Baras AS, De Marzo AM, Lotan TL. ERG and PTEN status of isolated high-grade PIN occurring in cystoprostatectomy specimens without invasive prostatic adenocarcinoma. *Hum Pathol.* 2016;55:117-125. doi:10.1016/j.humpath.2016.04.017.
29. Hieronymus H, Schultz N, Gopalan A, et al. Copy number alteration burden predicts prostate cancer relapse. *Proc Natl Acad Sci U S A.* 2014;111(30):11139-11144. doi:10.1073/pnas.1411446111.
30. Rubin MA, Girelli G, Demichelis F. Genomic Correlates to the Newly Proposed Grading Prognostic Groups for Prostate Cancer. *Eur Urol.* 2016;69(4):557-560. doi:10.1016/j.eururo.2015.10.040.
31. Berger MF, Lawrence MS, Demichelis F, et al. The genomic complexity of primary human prostate cancer. *Nature.* 2011;470(7333):214-220. doi:10.1038/nature09744.
32. Sun R, Hu Z, Sottoriva A, et al. Between-region genetic divergence reflects the mode and tempo of tumor evolution. *Nat Genet.* 2017;49(7):1015-1024. doi:10.1038/ng.3891.
33. McGranahan N, Swanton C. Biological and Therapeutic Impact of Intratumor Heterogeneity in Cancer Evolution. *Cancer Cell.* 2015;27(1):15-26. doi:10.1016/j.ccell.2014.12.001.
34. Landau DA, Carter SL, Stojanov P, et al. Evolution and Impact of Subclonal Mutations in Chronic Lymphocytic Leukemia. *Cell.* 2013;152(4):714-726. doi:10.1016/j.cell.2013.01.019.
35. Mroz EA, Rocco JW. MATH, a novel measure of intratumor genetic heterogeneity, is high in poor-outcome classes of head and neck squamous cell carcinoma. *Oral Oncol.* 2013;49(3):211-215. doi:10.1016/j.oraloncology.2012.09.007.
36. Sweeney CJ, Chen Y-H, Carducci M, et al. Chemohormonal Therapy in Metastatic Hormone-Sensitive Prostate Cancer. *N Engl J Med.* 2015;373(8):737-746. doi:10.1056/NEJMoa1503747.
37. Fizazi K, Tran N, Fein L, et al. Abiraterone plus Prednisone in Metastatic, Castration-Sensitive Prostate Cancer. *N Engl J Med.* June 2017. doi:10.1056/NEJMoa1704174.
38. James ND, Sydes MR, Clarke NW, et al. Addition of docetaxel, zoledronic acid, or both to first-line long-term hormone therapy in prostate cancer (STAMPEDE): survival results

- from an adaptive, multiarm, multistage, platform randomised controlled trial. *Lancet Lond Engl*. 2016;387(10024):1163-1177. doi:10.1016/S0140-6736(15)01037-5.
39. James ND, de Bono JS, Spears MR, et al. Abiraterone for Prostate Cancer Not Previously Treated with Hormone Therapy. *N Engl J Med*. 2017;0(0):null. doi:10.1056/NEJMoa1702900.
 40. Sowalsky AG, Ye H, Bubley GJ, Balk SP. Clonal progression of prostate cancers from Gleason grade 3 to grade 4. *Cancer Res*. 2013;73(3):1050-1055. doi:10.1158/0008-5472.CAN-12-2799.
 41. Wei L, Wang J, Lampert E, et al. Intratumoral and Intertumoral Genomic Heterogeneity of Multifocal Localized Prostate Cancer Impacts Molecular Classifications and Genomic Prognosticators. *Eur Urol*. July 2016. doi:10.1016/j.eururo.2016.07.008.
 42. Hong MKH, Macintyre G, Wedge DC, et al. Tracking the origins and drivers of subclonal metastatic expansion in prostate cancer. *Nat Commun*. 2015;6:6605. doi:10.1038/ncomms7605.
 43. Gudem G, Van Loo P, Kremeyer B, et al. The evolutionary history of lethal metastatic prostate cancer. *Nature*. 2015;520(7547):353-357. doi:10.1038/nature14347.
 44. Roubaud G, Liaw BC, Oh WK, Mulholland DJ. Strategies to avoid treatment-induced lineage crisis in advanced prostate cancer. *Nat Rev Clin Oncol*. 2017;14(5):269-283. doi:10.1038/nrclinonc.2016.181.

Figure legends

Figure 1. Intrafocal heterogeneity of index foci. A. Exome sequencing was performed on five to nine regions of a single contiguous index focus from ten cases of treatment-naïve prostate cancer. Presence (green) or absence (white) of mutations within each region sampled is shown. Regions are represented by columns, mutations by rows. Five locally advanced, node positive (LN pos) cases are above, five localized (LN neg) cases are below. Previously identified prostate cancer driver genes are indicated in black. Upper panels, the level of the radical prostatectomy (RP) specimen from which the region was sampled, from white (apex) to black (base), with purple (seminal vesicles) and red (lymph node metastasis) indicating extraprostatic regions. Left panels specify mutations as trunk (found in all regions), branch (at least two regions and not all regions), or leaf (identified in one region) mutations. B. Log2 plots of copy number estimates based on low coverage whole genome sequencing, except case 1009 based on exome sequencing. Left panel, the level of the prostatectomy specimen from which the region was sampled, as in A. Right panel, percent genome altered by copy number variation. C. Mean fraction of the genome altered in locally advanced (LN pos) and localized (LN neg) cases. Error bars represent standard deviation.

Figure 2. Correlation between genomic and histologic features. Network analysis of case 1010 based on shared mutations among eight regions of a contiguous index focus and one region from a lymph node metastasis is overlaid on a cartoon representation of the prostatectomy specimen. Corresponding H&E images of each region that underwent exome sequencing are depicted on the left. Met = lymph node metastasis.

Figure 3. Heterogeneity in CRPC-related alterations and predictive biomarkers. A.

Alterations across all cases in a panel of clinically relevant genes found to be altered in advanced

prostate cancer ⁶. B. Kaplan-meier analysis of biochemical recurrence between cases with

greater or fewer than the median CRPC-associated alterations (6), p-value <0.05 (log-rank test).

C. Fraction of intraprostatic regions within a case that harbor deletion and mutation or deep

deletion in five predictive biomarker genes, fraction of regions represented by size of the blue

circle (deletion) or “M” (mutation). Two cases (1010, 1022) had lymph node metastases

evaluated, represented by gray shading of the square.

Figure 4. Dispersion of subclones. A. Kernal density plot of variant allele frequency (VAF) of

trunk (brown), branch (orange), and leaf (yellow) mutations across all index focus mutations.

B,C. Calculated average number of mutations per diploid genome, or cancer cell fraction, for

each mutation identified in representative pairs of coincident regions. Each point represents one

mutation (B, case 1009; C, case 1010). Colors are as in A. D. VAF, relative to the mean VAF of

trunk mutations, of early prostate cancer driver gene mutations for all regions of the case. Gray,

diploid. Blue, one copy deleted.

Table 1
Clinical information for cases evaluated

	1009	1010	1015	1022	1029	1003	1020	1024	1027	1034
Stage category	Loc adv	Loc adv	Loc adv	Loc adv	Loc adv	Localized	Localized	Localized	Localized	Localized
Age at surgery	57	66	57	61	51	56	62	60	60	64
Ethnicity	AA	White	AA	White	White	White	White	White	White	AA
PSA (ng/mL)	8.35	12.76	43.05	NA	7.96	4.91	4.30	14.20	9.20	5.00
Pathologic stage	T3bN1	T3bN1	T3bN1	T3bN1	T3aN1	T2cN0	T2c	T3aN0	T3bN0	T2aN0
Gleason grade	4+5	4+5	3+4 (5)	4+5	4+3	4+5	3+4	3+4	4+3 (5)	4+3
Tumor volume (%)	40	60	65	45	25	15	25	25	25	20
Surgical margin	positive	negative	positive	positive	negative	negative	negative	positive	negative	negative
Nodes examined	22	20	18	14	15	5	4	22	22	13
Nodes involved	6	1	2	4+5	1	0	0	0	0	0
Germline sample	blood	blood	blood	blood	blood	blood	blood	blood	benign prostate	blood
Prostate or SV regions	7	8	8	6	8	5	9	8	7	6
Lymph node regions	0	1	0	1	0	0	0	0	0	0
Time to recurrence (mos)	1	2	3	NA	NA	NA	NA	NA	NA	NA
Additional therapy	Yes	Yes	Yes	Yes	No	No	No	No	No	No

Figure 1

[Click here to download Figure Figure 1. Oncoprints bw 15.png](#)

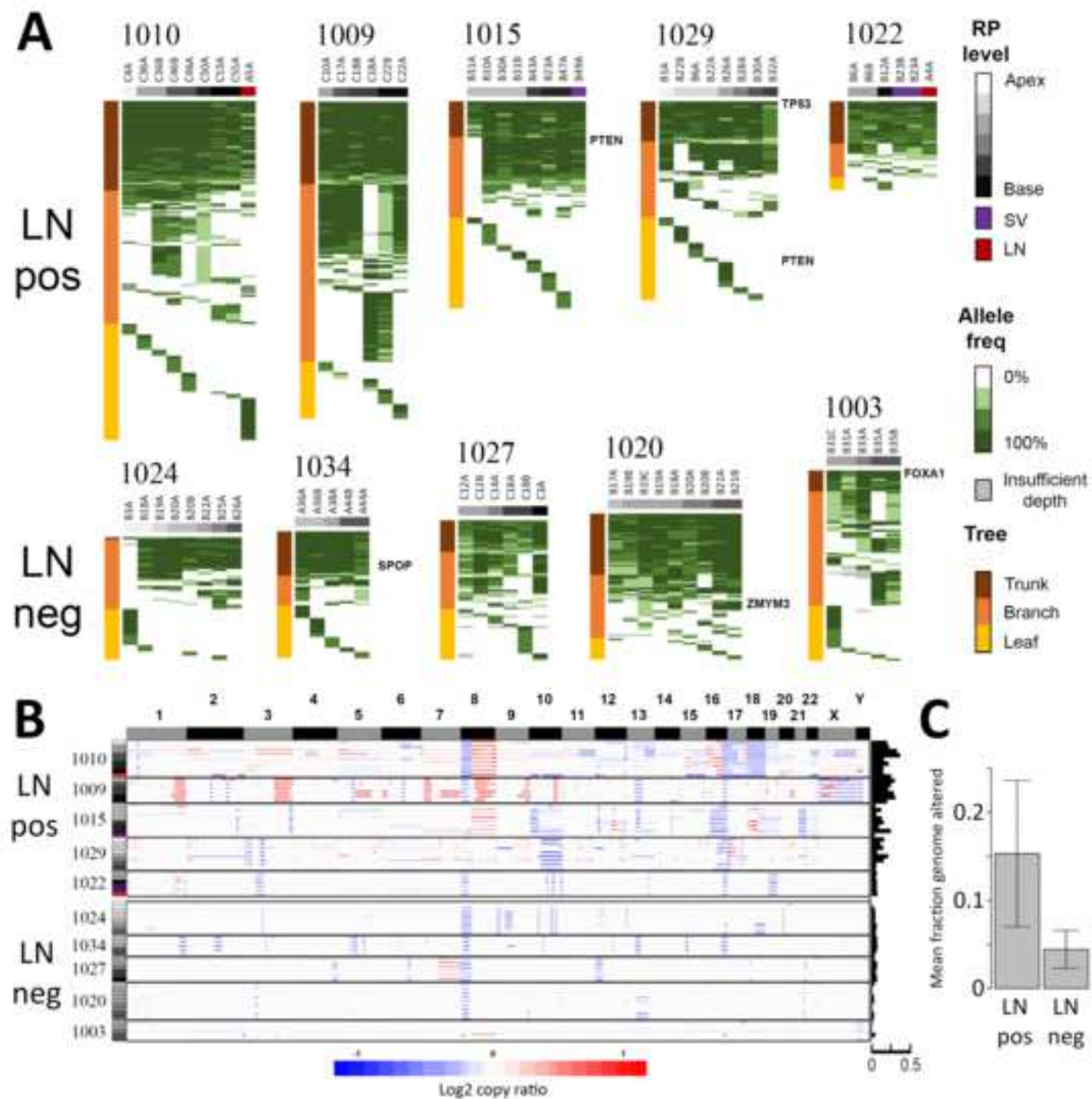


Figure 2

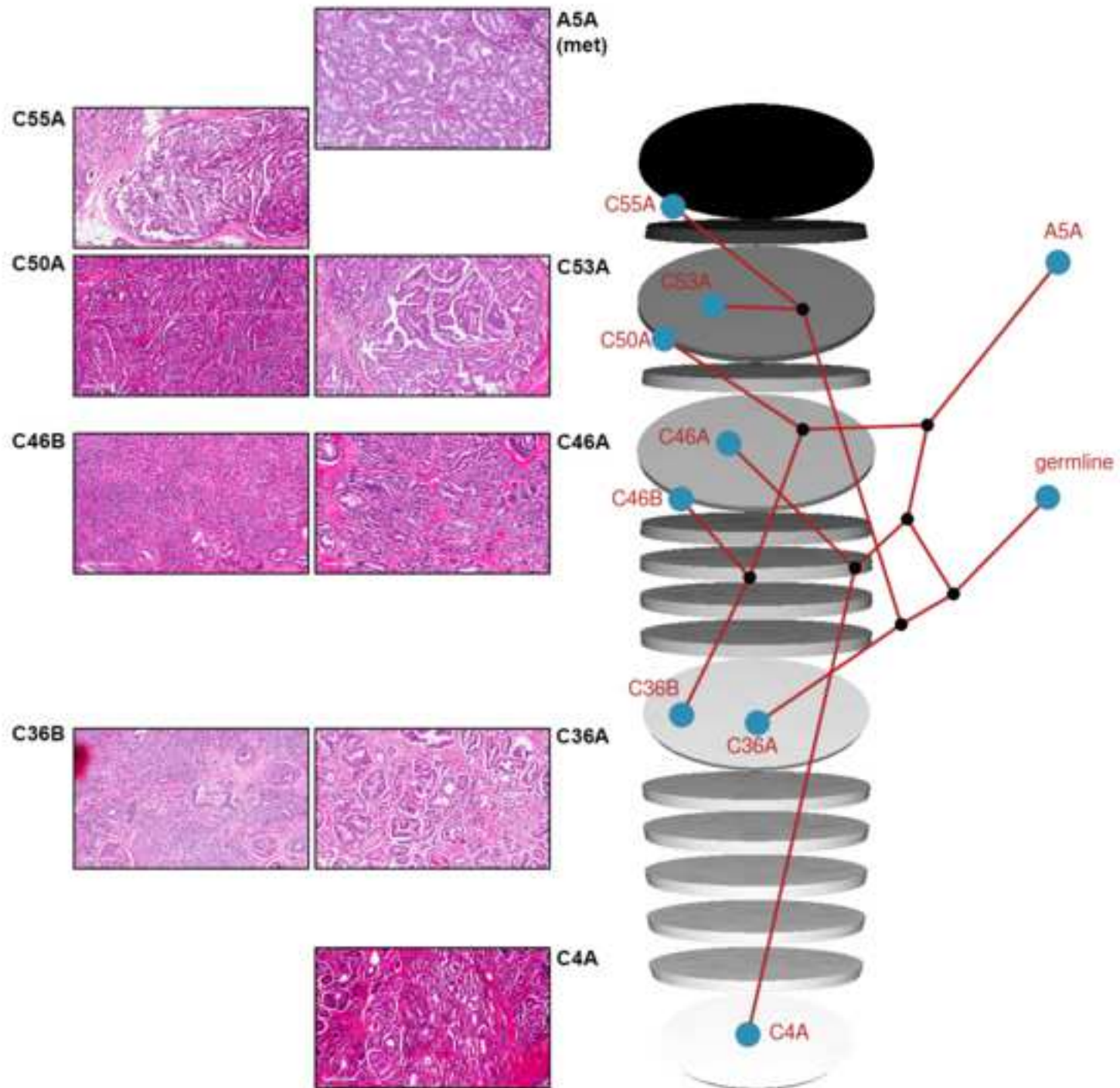


Figure 3

Click here to download Figure Fig 3 SU2C drivers image with Nexus and drugs v9.png

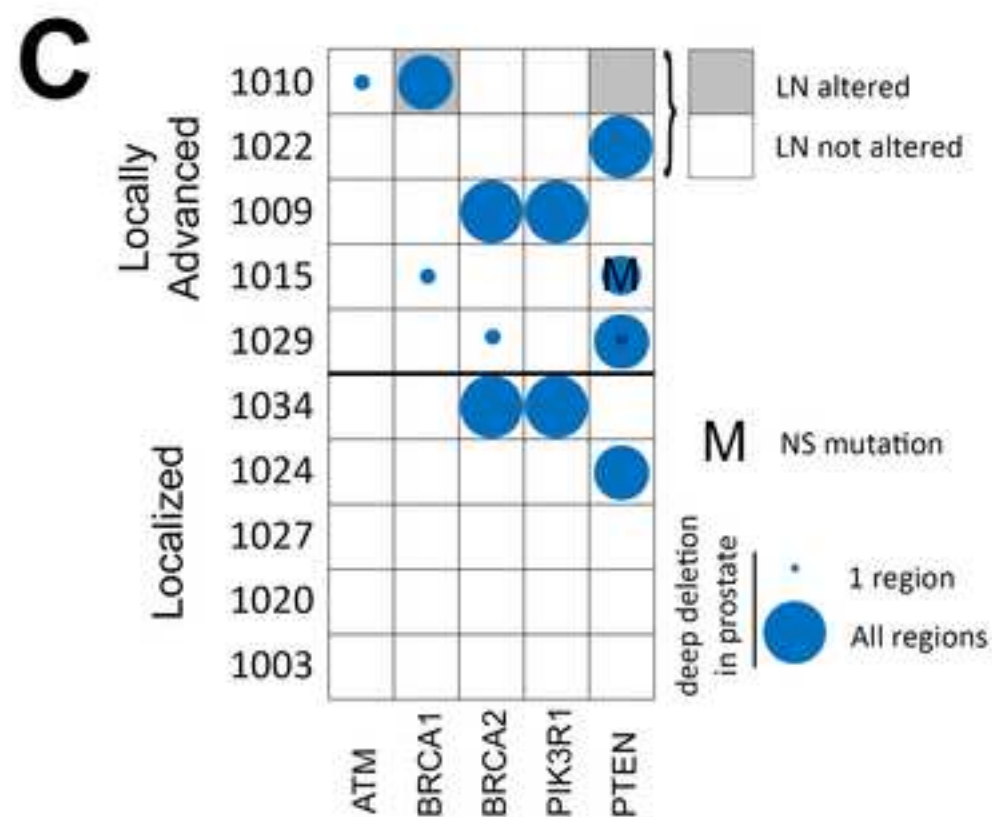
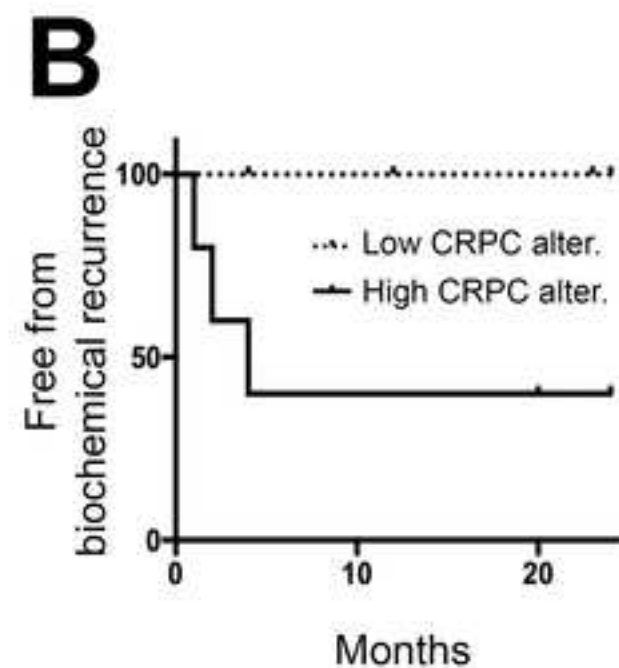
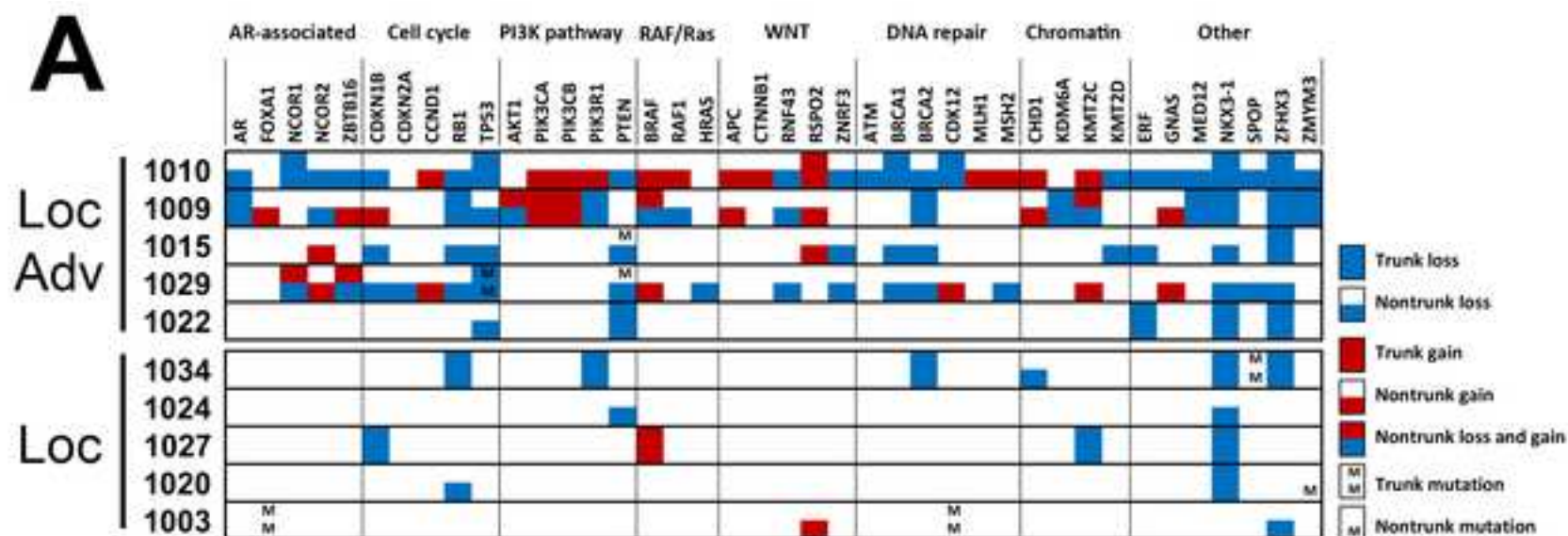
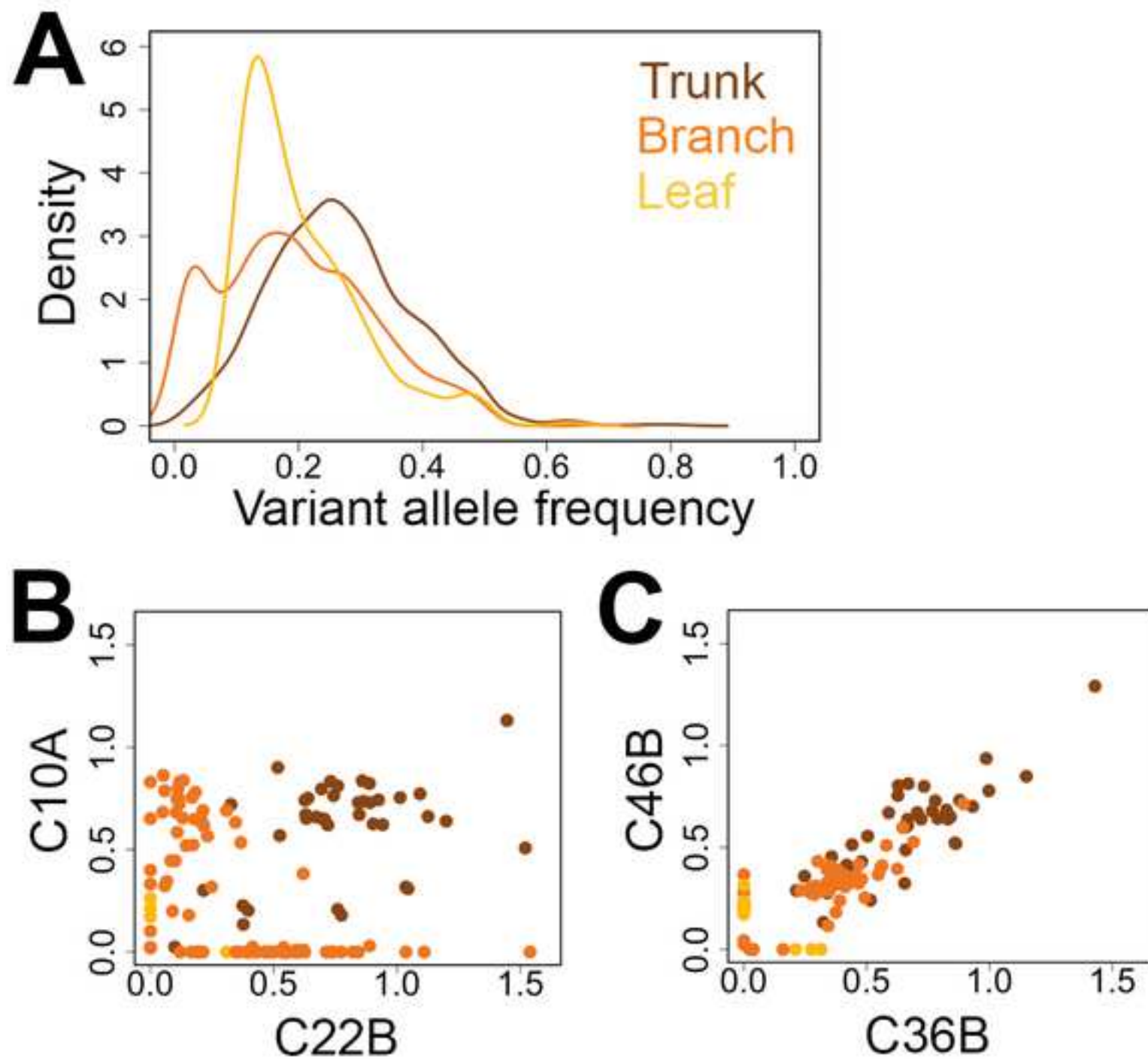


Figure 4

[Click here to download Figure Figure 4 Subclone mixing v2.png](#)





[Click here to access/download](#)

Supplemental Material, Other

Multiregion seq manuscript supplemental methods Final
v2.docx

